VECTOR-VALUED STATISTICS OF BINOMIAL PROCESSES: BERRY–ESSEEN BOUNDS IN THE CONVEX DISTANCE

BY MIKOŁAJ J. KASPRZAK[a] AND GIOVANNI PECCATI[b]

Department of Mathematics, University of Luxembourg, [a]*mikolaj.kasprzak@uni.lu,* [b]*giovanni.peccati@uni.lu*

We study the discrepancy between the distribution of a vector-valued functional of i.i.d. random elements and that of a Gaussian vector. Our main contribution is an explicit bound on the convex distance between the two distributions, holding in every dimension. Such a finding constitutes a substantial extension of the one-dimensional bounds deduced in Chatterjee (*Ann. Probab.* **36** (2008) 1584–1610) and Lachièze-Rey and Peccati (*Ann. Appl. Probab.* **27** (2017) 1992–2031), as well as of the multidimensional bounds for smooth test functions and indicators of rectangles derived, respectively, in Dung (*Acta Math. Hungar.* **158** (2019) 173–201), and Fang and Koike (*Ann. Appl. Probab.* **31** (2021) 1660–1686). Our techniques involve the use of Stein's method, combined with a suitable adaptation of the recursive approach inaugurated by Schulte and Yukich (*Electron. J. Probab.* **24** (2019) 1–42): this yields rates of converge that have a presumably optimal dependence on the sample size. We develop several applications of a geometric nature, among which is a new collection of multidimensional quantitative limit theorems for the intrinsic volumes associated with coverage processes in Euclidean spaces.

1. Introduction.

1.1. *Overview.* Let $X = (X_1, \ldots, X_n)$ be a vector of independent random elements with values in a measurable space $(\mathcal{X}, \mathbb{X})$, and let $f : \mathcal{X}^n \to \mathbb{R}^d$, $d \geq 1$, be a measurable mapping. The aim of this paper is to extend the techniques introduced in [6, 24] (in the case $d = 1$) in order to provide explicit bounds on the distance between the distribution of the d-dimensional random vector $f(X)$, and that of a suitable Gaussian element. Our principal contribution is an explicit bound on the *convex distance* between the two distributions (see Theorem 2.2 below), holding for arbitrary values of d. As discussed in Section 2.3 below, such findings substantially refine and complement the recent estimates from [13, 15], where bounds on the normal approximation of $f(X)$ are obtained for smooth test functions of class C^2 and C^3 (see [13], Theorems 3.1 and 3.2, as well as Theorem 2.1 of the present paper), and for indicators of hyperrectangles (see [15], Theorem 1.3).

Our main results are based on the use of the multidimensional *Stein's method* for normal approximations (see e.g., [10], Chapter 12, and [33], Chapter 4), which is combined with a recursive method devised for dealing with indicators of convex sets, inaugurated in [45] in the context of geometric functionals of Poisson random measures—see [25, 35, 46] for further applications of this methodology. An important point (see again Section 2.3) is that the use of the approach of [45] allows one to remove from the bounds all spurious logarithmic dependencies in the parameter n (sample size), such as those appearing for example, in the main bounds of [15, 41]. Further references exploiting Stein's method in the context of multivariate normal approximations are [1, 8, 14, 16, 20, 29, 39–42, 47].

We demonstrate the flexibility and scope of our results by developing two applications:

(i) in Section 3, to a multivariate central limit theorem (CLT) for the intrinsic volumes associated with the Euclidean Boolean model over increasing domains—thus extending the findings of [6, 19, 24] (which contain one-dimensional CLTs Euclidean volumes) and [22, 45] (where multidimensonal CLTs are proved in the context of the Poisson-based Boolean model);

(ii) in Section 4, to multivariate quantitative CLTs for statistics depending on k-nearest neighbour graphs, thus recovering a multivariate version of central results from [4, 6].

1.2. *Structure of the paper.* Section 1.3 discusses the notational conventions adopted throughout the paper. Our main abstract results are stated and discussed in Section 2. Section 3 and Section 4 are devoted, respectively, to our main findings for coverage processes and for structures with local dependence. The proofs of our abstract results are presented in full detail in Section 5, whereas Section 6 deals respectively, with the proofs of our results for intrinsic volumes of Boolean models. The supplementary material [23] contains further technical details of the arguments presented in Section 6, as well as all the proofs related to quantitative CLTs for functionals of structures with local dependence.

1.3. *Setup and notation.* *Standard conventions.* For any $d \in \mathbb{N}$, $\mathbb{R}^d$ is always equipped with the Euclidean topology, the brackets $\langle \cdot, \cdot \rangle$ denote the Euclidean inner product and $\| \cdot \|$ stands for the Euclidean norm (the dependence on the dimension is removed to simplify the notation). The symbol $\langle \cdot, \cdot \rangle_{\text{H.S.}}$ indicates the Hilbert–Schmidt inner product, $\| \cdot \|_{\text{H.S.}}$ is the Hilbert–Schmidt norm and $\| \cdot \|_{\text{op}}$ is the (Euclidean) operator norm. Given $d \in \mathbb{N}$ and a mapping $F : \mathbb{R}^d \to \mathbb{R}$ sufficiently smooth, we write $\text{Hess}(F)(x)$ for the Hessian of F at $x \in \mathbb{R}^d$ and, for $k \geq 0$, we use the symbol $D^k F$ to indicate the kth (Fréchet) derivative of F, that is: $D^k F(x)$ stands for the k-linear form on $\mathbb{R}^d$ given by

$$D^k F(x)[u_1, \ldots, u_k] := \sum_{i_1, \ldots, i_k = 1}^{d} \frac{\partial^k F}{\partial x_{i_1} \cdots \partial x_{i_k}} (x)(u_1)_{i_1} \cdots (u_k)_{i_k},$$

for $x, u_1, \ldots, u_k \in \mathbb{R}^d$ and $(u_i)_j$ denoting the jth component of the vector u_i. We also set

$$(1.1) \qquad M_k(F) := \sup_{x \in \mathbb{R}^d} \| D^k F(x) \|_{\text{op}}, \qquad \tilde{M}_2(F) := \sup_{x \in \mathbb{R}^d} \| \text{Hess}(F)(x) \|_{\text{H.S.}}.$$

Objects such as $M_k(F)$ and $\tilde{M}_2(F)$ (and variations thereof) play an important role in several applications of the multidimensional Stein's method—see [8, 11, 29, 39] for a sample. Every random object appearing below is assumed to be defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with $\mathbb{E}$ denoting expectation with respect to $\mathbb{P}$.

Convex distance. Given $d \in \mathbb{N}$ and random vectors A, B with values in $\mathbb{R}^d$, we define the *convex distance* between the distributions of A and B as

$$(1.2) \qquad d_{\text{convex}}(A, B) := \sup_{h \in \mathcal{I}_d} |\mathbb{E}h(A) - \mathbb{E}h(B)|,$$

where $\mathcal{I}_d$ is the set of all indicator functions of measurable convex subsets of $\mathbb{R}^d$. For $n \in \mathbb{N}$, we also write $[n] := \{1, \ldots, n\}$.

Difference operators. In this paper, we will extensively use some discrete operators defined as the difference between a given function of $X = (X_1, \ldots, X_n)$ (or of some perturbation of it) and the same function computed on some partial resampling of X. These operators naturally emerge in functional inequalities connected to Efron–Stein estimates (see e.g., [5], Chapters 3–5), and are pivotal objects in the theory developed in [6, 13, 15, 24] and its

extensions and applications (mainly via the use of the quantities T_A defined in (1.3) below—see, for instance, [9, 12, 18]). Now fix $n \in \mathbb{N}$, let $(\mathcal{X}, \mathbb{X})$ be a measurable space and let $X_1, X_2, \ldots$ be a sequence of independent $\mathcal{X}$-valued random elements. Let $X = (X_1, \ldots, X_n)$ and write $X' = (X'_1, \ldots, X'_n)$ and $\tilde{X} = (\tilde{X}_1, \ldots, \tilde{X}_n)$ to indicate copies of X, such that X, X', $\tilde{X}$ are mutually independent. For each $A \subset [n]$, we define

$$X_i^A = \begin{cases} X'_i & \text{if } i \in A, \\ X_i & \text{if } i \notin A. \end{cases}$$

For $j \in [n]$, we write X^j for $X^{\{j\}}$ and, for any $d \in \mathbb{N}$ and $f : \mathcal{X}^n \to \mathbb{R}^d$, we set

$$\Delta_j f(X) = f(X) - f(X^j).$$

Given $A \subset [n]$, we introduce the notation

(1.3)
$$T_A := \sum_{j \notin A} [\Delta_j f(X)][\Delta_j f(X^A)]^T \in \mathbb{R}^{d \times d},$$

and

$$T := \frac{1}{2} \sum_{A \subsetneq [n]} k_{n,A} T_A,$$

where

(1.4)
$$k_{n,A} := \frac{1}{\binom{n}{|A|}(n - |A|)}, \quad A \subsetneq [n].$$

Finally, for any $d \in \mathbb{N}$ and function $g : \mathcal{X}^{2n} \to \mathbb{R}^d$, we set the following notation:

$$\tilde{\Delta}_i g(X, X') := g(X, X') - g((X_1, \ldots, X_{i-1}, \tilde{X}_i, X_{i+1}, \ldots, X_n), X').$$

The following statement from [6] is exploited in several parts of the paper (see e.g., Remark 2.3 below).

LEMMA 1.1 (Lemma 2.3 in [6]). *For any $g, h : \mathcal{X}^n \to \mathbb{R}$ such that $\mathbb{E}g(X)^2$ and $\mathbb{E}h(X)^2$ are both finite, we have*

$$\text{Cov}(g(X), h(X)) = \frac{1}{2} \sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \mathbb{E}[\Delta_j g(X) \Delta_j h(X^A)],$$

where $k_{n,A}$ is defined in (1.4) and Cov *indicates the covariance.*

2. Main results. Fix $d \geq 1$ and let $f : \mathcal{X}^n \to \mathbb{R}^d$ be a measurable function such that the random vector

$$W := f(X)$$

verifies the relations $\mathbb{E}W = 0$ and $\mathbb{E}\|W\|^2 < \infty$. Let $\Sigma \in \mathbb{R}^{d \times d}$ be nonnegative definite and let N_Σ be d-dimensional vector with mean zero and covariance Σ. We will now present the main results of the paper: detailed comparisons with the existing literature are gathered together in Section 2.3, whereas proofs can be found in Section 5.

2.1. *General estimates.*   We start with a collection of inequalities involving smooth test functions: they are both a multidimensional version of [6], Theorem 2.2, and an alternative to the bounds appearing in [13], Theorems 3.1 and 3.2.

THEOREM 2.1.   *For any* $F \in C^3(\mathbb{R}^d)$,

$$(2.1) \quad |\mathbb{E}F(W) - \mathbb{E}F(N_\Sigma)| \leq \frac{\tilde{M}_2(F)}{2} \mathbb{E}\|\mathbb{E}[T|X] - \Sigma\|_{\text{H.S.}} + \frac{M_3(F)}{12} \sum_{j=1}^{n} \mathbb{E}\|\Delta_j f(X)\|^3.$$

Moreover, if Σ is positive definite, then, for any $F \in C^2(\mathbb{R}^d)$,

$$|\mathbb{E}F(W) - \mathbb{E}F(N_\Sigma)| \leq M_1(F)\|\Sigma^{-1}\|_{\text{op}}\mathbb{E}\|\mathbb{E}[T|X] - \Sigma\|_{\text{H.S.}}$$

$$(2.2) \qquad\qquad\qquad + \frac{M_2(F)\sqrt{2\pi}}{16}\|\Sigma^{-1}\|_{\text{op}} \sum_{j=1}^{n} \mathbb{E}\|\Delta_j f(X)\|^3.$$

We now move on to a bound on the convex distance between the distribution of W and that of a centered Gaussian random vector N_Σ such that Σ is invertible. Our estimate is valid under the additional assumption that

$$(2.3) \qquad\qquad\qquad \mathbb{E}\|\Delta_j f(X)\|^6 < \infty, \quad j \in [n],$$

and is expressed in terms of the parameters γ_1, γ_2, γ_3 and γ_4 defined as follows:

$$\gamma_1 := \sum_{j=1}^{n} \mathbb{E}\|\Delta_j f(X)\|^3$$

$$\gamma_2 := \left(\sum_{j=1}^{n} \mathbb{E}\|\Delta_j f(X)\|^4\right)^{1/2}$$

$$\gamma_3 := \left\{\frac{3}{2}\sum_{i=1}^{n}\mathbb{E}\left[\left(\sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \mathbb{1}_{[\tilde{\Delta}_i\Delta_j f(X) \neq 0]}\sqrt{\|\Delta_j f(X)\| + \|\tilde{\Delta}_i\Delta_j f(X)\|}\right.\right.\right.$$

$$\left.\left. \cdot \|\Delta_j f(X^A)\|\|\Delta_j f(X)\|\right)^2\right]$$

$$+ 9\sum_{i=1}^{n}\mathbb{E}\left[\left(\sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \sqrt{\|\Delta_j f(X)\| + \|\tilde{\Delta}_i\Delta_j f(X)\|}\|\tilde{\Delta}_i\Delta_j f(X^A)\|\|\Delta_j f(X)\|\right)^2\right]$$

$$+ 9\sum_{i=1}^{n}\mathbb{E}\left[\left(\sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \sqrt{\|\Delta_j f(X)\| + \|\tilde{\Delta}_i\Delta_j f(X)\|}\|\Delta_j f(X^A)\|\|\tilde{\Delta}_i\Delta_j f(X)\|\right)^2\right]$$

$$+ 9\sum_{i=1}^{n}\mathbb{E}\left[\left(\sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \sqrt{\|\Delta_j f(X)\| + \|\tilde{\Delta}_i\Delta_j f(X)\|}\right.\right.$$

$$\left.\left.\left. \cdot \|\tilde{\Delta}_i\Delta_j f(X^A)\|\|\tilde{\Delta}_i\Delta_j f(X)\|\right)^2\right]\right\}^{1/3},$$

$$\gamma_4 := \left\{\frac{3}{2}\sum_{i=1}^{n}\mathbb{E}\left[\left(\sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \mathbb{1}_{[\tilde{\Delta}_i\Delta_j f(X) \neq 0]}\sqrt{\|\Delta_j f(X)\|^2 + \|\tilde{\Delta}_i\Delta_j f(X)\|^2}\right.\right.\right.$$

$$\left.\left. \cdot \|\Delta_j f(X^A)\|\|\Delta_j f(X)\|\right)^2\right]$$

$$+ \frac{27}{4} \sum_{i=1}^{n} \mathbb{E}\Bigg[\Bigg(\sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \sqrt{\|\Delta_j f(X)\|^2 + \|\tilde{\Delta}_i \Delta_j f(X)\|^2}$$

$$\cdot \|\tilde{\Delta}_i \Delta_j f(X^A)\| \|\Delta_j f(X)\|\Bigg)^2\Bigg]$$

$$+ \frac{27}{4} \sum_{i=1}^{n} \mathbb{E}\Bigg[\Bigg(\sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \sqrt{\|\Delta_j f(X)\|^2 + \|\tilde{\Delta}_i \Delta_j f(X)\|^2}$$

$$\cdot \|\Delta_j f(X^A)\| \|\tilde{\Delta}_i \Delta_j f(X)\|\Bigg)^2\Bigg]$$

$$+ \frac{27}{4} \sum_{i=1}^{n} \mathbb{E}\Bigg[\Bigg(\sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \sqrt{\|\Delta_j f(X)\|^2 + \|\tilde{\Delta}_i \Delta_j f(X)\|^2}$$

$$\cdot \|\tilde{\Delta}_i \Delta_j f(X^A)\| \|\tilde{\Delta}_i \Delta_j f(X)\|\Bigg)^2\Bigg]\Bigg\}^{1/4},$$

where we have used the notation (1.4).

The next statement is the main contribution of our work.

THEOREM 2.2. *Let $\Sigma \in \mathbb{R}^{d \times d}$ be positive-definite, suppose that* (2.3) *is verified and let*

$$\gamma := \max\Big\{\sqrt{\mathbb{E}\|\mathbb{E}[T - \Sigma|X]\|_{\text{H.S.}}^2}, \gamma_1, \gamma_2, \gamma_3, \gamma_4\Big\}.$$

Then,

(2.4) $$d_{\text{convex}}(W, N_\Sigma) \leq 541 d^4 \max\{1, \|\Sigma^{-1}\|_{\text{op}}^2\}\gamma.$$

REMARK 2.3. If Σ is the covariance matrix of W, then $\Sigma = \mathbb{E}T$ (by virtue of Lemma 1.1) and

$$\mathbb{E}\|\mathbb{E}[T - \Sigma|X]\|_{\text{H.S.}} \leq \sqrt{\mathbb{E}\|\mathbb{E}[T - \Sigma|X]\|_{\text{H.S.}}^2} = \sqrt{\sum_{k,l=1}^{d} \text{Var}[\mathbb{E}[T_{k,l}|X]]}$$

and the bounds in Theorems 2.1 and 2.2 are simplified accordingly.

2.2. *A simplified bound for symmetric statistics.* We will now present a useful statement (Lemma 2.4 below), allowing one to bound the quantity $\mathbb{E}\|\mathbb{E}[T - \mathbb{E}T|X]\|_{\text{H.S.}}^2$ whenever f is symmetric and the random elements X_i are (independent and) identically distributed. We need some additional notation: denoting by $\{X_i'\}$ an independent copy of $\{X_i\}$, for any random vector $Z = (Z_1, \ldots, Z_n)$ and every $A \subset [n]$, we set

$$Z_i^A = \begin{cases} X_i' & \text{if } i \in A, \\ Z_i & \text{if } i \notin A \end{cases} \quad \text{and} \quad Z^A = (Z_1^A, \ldots Z_n^A);$$

also, for $1 \leq i \neq j \leq n$, we write

$$\Delta_i f(Z) := f(Z) - f(Z^{\{i\}}); \qquad \Delta_{i,j} f(Z) := f(Z) - f(Z^{\{i\}}) - f(Z^{\{j\}}) + f(Z^{\{i,j\}}).$$

We say that $Z = (Z_1, \ldots, Z_n)$ is a *recombination* of $\{X, X', \tilde{X}\}$ if $Z_i \in \{X_i, X_i', \tilde{X}_i\}$ for all $i \in [n]$. Finally, we let

$$B_n(f) := \sup_{(Y,Z,Z')} \mathbb{E}[\mathbb{1}_{\{\Delta_{1,2}f(Y) \neq 0\}}\|\Delta_1 f(Z)\|^2 \|\Delta_2 f(Z')\|^2];$$

$$B'_n(f) := \sup_{(Y,Y',Z,Z')} \mathbb{E}\big[\mathbb{1}_{\{\Delta_{1,2}f(Y)\neq 0, \Delta_{1,3}f(Y')\neq 0\}} \|\Delta_2 f(Z)\|^2 \|\Delta_3 f(Z')\|^2\big],$$

where the suprema run over all vectors Y, Y', Z, Z' that are recombinations of $\{X, X', \tilde{X}\}$. The proof of the following result can be found in Section 5.3.

LEMMA 2.4 (cf. Theorem 5.1 of [24]). *Suppose that f is symmetric and that $X_1, \dots, X_n$ are i.i.d. Then,*

$$\sqrt{\mathbb{E}\|\mathbb{E}[T|X] - \mathbb{E}T\|_{\text{H.S.}}^2} \leq 4\sqrt{n}\Big(\sqrt{nB_n(f)} + \sqrt{n^2 B'_n(f)} + \sqrt{\mathbb{E}\|\Delta_1 f(X)\|^4}\Big).$$

EXAMPLE 1. Fix $d \geq 1$ and let $X_1, \dots, X_n$ be i.i.d. $\mathbb{R}^d$-valued random variables with mean zero, covariance matrix Σ and satisfying $\mathbb{E}\|X_1\|^4 < \infty$. Let $W = f(X) = \frac{1}{\sqrt{n}}\sum_{i=1}^n X_i$. Adopt the notation of Theorem 2.2 and note that

(2.5)
$$\gamma_1 = \frac{1}{n^{3/2}} \sum_{j=1}^n \mathbb{E}\|X_j - X'_j\|^3 \leq 8\mathbb{E}\|X_1\|^3 n^{-1/2};$$

$$\gamma_2 = \frac{1}{n}\Big(\sum_{j=1}^n \mathbb{E}\|X_j - X'_j\|^4\Big)^{1/2} \leq 4\sqrt{\mathbb{E}\|X_1\|^4} n^{-1/2}.$$

Moreover, note that $\tilde{\Delta}_i \Delta_j f(X) = \mathbb{1}_{[i=j]}(X_j - \tilde{X}_j)$ and that, for $A \subsetneq [n]$ and $j \notin A$, $\Delta_j(X^A) = \Delta_j(X)$. It follows that, if $\mathbb{E}\|X_1\|^6 < \infty$ then, for $p = 1, 2$,

(2.6)
$$\gamma_{p+2} = \Bigg\{ \frac{3}{2} \sum_{j=1}^n \mathbb{E}\big[(\|X_j - X'_j\|^p + \|X_j - \tilde{X}_j\|^p)\|X_j - X'_j\|^4\big] n^{-2-p/2}$$

$$+ 2\Big(\frac{9}{2} + \frac{9}{2p}\Big)$$

$$\cdot \sum_{j=1}^n \mathbb{E}\big[(\|X_j - X'_j\|^p + \|X_j - \tilde{X}_j\|^p)\|X_j - \tilde{X}_j\|^2 \|X_j - X'_j\|^2\big] n^{-2-p/2}$$

$$+ \Big(\frac{9}{2} + \frac{9}{2p}\Big)$$

$$\cdot \sum_{j=1}^n \mathbb{E}\big[(\|X_j - X'_j\|^p + \|X_j - \tilde{X}_j\|^p)\|X_j - \tilde{X}_j\|^4\big] n^{-2-p/2}\Bigg\}^{1/(p+2)}$$

$$\leq 2^{1+1/(p+2)}\Big(30 + \frac{27}{p}\Big)^{1/(p+2)} (\mathbb{E}\|X_1\|^{p+4})^{1/(p+2)} n^{-1/2}.$$

Furthermore, since f is symmetric, we can apply Lemma 2.4. Adopting the notation thereof and letting Y be a recombination of $\{X, X', \tilde{X}\}$, we have that

$$\Delta_{1,2}f(Y)$$

$$= \frac{1}{\sqrt{n}}\bigg[\Big(\sum_{i=1}^n Y_i\Big) - \Big(X'_1 + \sum_{i=2}^n Y_i\Big) - \Big(X'_2 + Y_1 + \sum_{i=3}^n Y_i\Big) + \Big(X'_1 + X'_2 + \sum_{i=3}^n Y_i\Big)\bigg] = 0.$$

Therefore, noting that $\Sigma = \mathbb{E}T$ (see Remark 2.3), we have that

(2.7)
$$\sqrt{\mathbb{E}\|\mathbb{E}[T|X] - \Sigma\|_{\text{H.S.}}^2} \leq 16\sqrt{\mathbb{E}\|X_1\|^4} n^{-1/2}.$$

Finally, combining Theorems 2.1 and 2.2 with (2.5)–(2.7) and letting N_Σ be a centred normal random variable with covariance Σ, we obtain:

1. For $F \in C^3(\mathbb{R}^d)$,

$$(2.8) \qquad \left| \mathbb{E}[F(W)] - \mathbb{E}[F(N_\Sigma)] \right| \le \left(8\tilde{M}_2(F)\sqrt{\mathbb{E}\|X_1\|^4} + \frac{3M_3(F)}{4}\mathbb{E}\|X_1\|^3 \right)n^{-1/2};$$

2. If Σ is positive definite then, for $F \in C^2(\mathbb{R}^d)$,

$$(2.9) \qquad \begin{aligned} &\left| \mathbb{E}[F(W)] - \mathbb{E}[F(N_\Sigma)] \right| \\ &\le \|\Sigma^{-1}\|_{\mathrm{op}}\left(16M_1(F)\sqrt{\mathbb{E}\|X_1\|^4} + \frac{M_2(F)\sqrt{2\pi}}{2}\mathbb{E}\|X_1\|^3 \right)n^{-1/2}; \end{aligned}$$

3. If Σ is positive definite and $\mathbb{E}\|X_1\|^6 < \infty$ then

$$(2.10) \qquad \begin{aligned} d_{\mathrm{convex}}(W, N_\Sigma) &\le 8656d^4 \max\{1, \|\Sigma^{-1}\|_{\mathrm{op}}^2\} \\ &\quad \cdot \max\{\sqrt{\mathbb{E}\|X_1\|^4}, \mathbb{E}\|X_1\|^3, (\mathbb{E}\|X_1\|^5)^{1/3}, (\mathbb{E}\|X_1\|^6)^{1/4}\}n^{-1/2}. \end{aligned}$$

As a consequence, in the three cases our estimates yield bounds on normal approximations converging to zero with an optimal rate which is commensurate to the square-root of the sample size.

2.3. *Literature review.* As already discussed in the Introduction, our main theoretical bounds (2.1), (2.2) and (2.4) generalise the one-dimensional estimates stated in [7], Theorem 2.2 (1-dimensional Wasserstein distance) and [24], Theorem 4.2 (1-dimensional Kolmogorov distance). One noticeable difference between (2.4) and the content of [24], Theorem 4.2, is that the latter allows one to consider random variables such that $\mathbb{E}|\Delta_j f(X)|^4 < \infty$ (which is weaker than (2.3) above). One could also similarly relax the moment assumptions in our bounds, at the cost of considerably longer proofs: since this point would impact only marginally the scope of our applications, we decided not to pursue it further. One can check that, in dimension $d = 1$, relation (2.4) yields upper bounds on the Kolmogorov distance that are commensurate to [24], equation (4.2) (which requires integrability assumptions analogous to (2.3)).

In the multidimensional case, the most relevant references related to our work are [13, 15]. The proof and statements of Theorems 3.1 and 3.2 in [13] deal with the case of smooth test functions, and are close to our estimates (2.1)–(2.2) (albeit expressed in terms of slightly different quantities). Although we did not carry out the computations in full detail, it is reasonable to expect that, in the case of smooth test functions, the upper bounds from [13] would yield rates of convergence similar to ours in all applications we will develop below.

Theorem 1.3 in [15] provides an upper bound on the quantity

$$\mathbf{D} := \sup_R \left| \mathbb{P}[W \in R] - \mathbb{P}[N_\Sigma \in R] \right|,$$

where R runs over the class of all rectangles of $\mathbb{R}^d$, in such a way that $\mathbf{D} \le d_{\mathrm{convex}}(W, N_\Sigma)$. The bounds on $\mathbf{D}$ derived in [15] are expressed in terms of the same difference operators appearing on the right-hand side of (2.4), and are obtained by combining Götze's "generator approach" towards Stein's method [20] with some smoothing estimates from [3]. Apart from the fact that the class of convex sets is strictly larger than that of rectangles, the main differences between [15], Theorem 1.3, and (2.4) are the following: (i) because of the involved techniques, the bounds from [15] tend to produce a suboptimal dependence on the sample size n, typically polynomially worse (see e.g., [15], Corollary 1.3) than that given by our

bounds; in comparison, our bounds have been devised in order to produce presumably optimal rates (i.e., rates proportional to the inverse of the square root of the variance of W, as $n \to \infty$); (ii) the upper bounds deduced in [15] typically display a logarithmic-type dependence on the dimension d, which is in stark contrast with the d^4 prefactor appearing in our inequalities; (iii) the estimates from [15] require weaker integrability assumptions on the random vectors $\Delta_j f(X)$ (typically, finite fourth moments vs. finite sixth moments in our case). The better performance of our bounds described at Point (i) is a direct consequence of the recursive approach initiated in [45], and further developed in the present work (see also [25, 35]). The discrepancy at Point (ii) is somehow consistent with the fundamental geometric fact that the Gaussian isoperimetric constant for convex sets increases as a power of d, whereas the one of rectangles is bounded from above by a multiple of $\sqrt{\log d}$ (see e.g., [32, 38], as well as [35], Appendix A), and is likely to be unavoidable; note that a similar dependence on the dimension (of the order of d^5) also appears in the main bounds developed in [45]. As already observed in the case of smooth test functions, one could in principle push further the approach developed in the present work (at the cost of considerably longer proofs) in order to relax the integrability requirements on the random vectors $\Delta_j f(X)$, thus addressing Point (iii).

It is also natural to compare the specific content of Example 1 with the best available Berry–Esseen bounds for the classical multidimensional CLT, see for example, [2, 14]. In this case, our estimate (2.10) yields an upper bound converging to zero at the optimal rate $n^{-1/2}$, displaying nonetheless a suboptimal dimensional dependence of the order d^4 (to be contrasted with Bentkus' famous bounds from [2], featuring a dimensional dependence of the order $d^{1/4}$), as well as requiring stronger integrability assumptions (i.e., existence of 6th moments). In principle, these shortcomings are a consequence of the fact that our methods are not optimised for dealing with the case of linear statistics, which often require careful ad-hoc arguments—see for example, [2, 14, 15]. We refer the reader to [15], Section 1.1, and the references therein, for a detailed discussion of the state-of-the-art literature on the matter.

3. Applications to coverage processes. We will now use our main estimates in order to assess the joint fluctuations of geometric quantities associated with a binomial version of *the (Euclidean) Boolean model* in the so-called *thermodynamic regime* (i.e., with grains of a fixed radius). Our main results are gathered together in the statements of Theorem 3.1 and Theorem 3.2 below, whose proofs are postponed to Section 6. As explained in Remark 3.4, our results are the first quantitative CLTs for geometric functionals attached to the binomial Boolean model, and should be compared with recent advances in a Poissonian setting [22, 45]. We refer the reader to the monographs [21, 36, 44], as well as [26], Chapters 16 and 17, for a general introduction to Boolean models and for an overview of the history of the subject. All needed geometric notions (in particular, related to convex geometry and intrinsic volumes) can be found in [44], Chapter 14.

Fix $d \geq 1$. For every $n \geq 1$, we write E_n to indicate a cube of volume n in $\mathbb{R}^d$ and let $C_1, \ldots, C_n$ denote i.i.d. random variables uniformly distributed on E_n (customarily called *germs*). Let $K = B(0, R) \subset \mathbb{R}^d$ be the closed ball of radius $R > 0$ centered at the origin and define $X_i := C_i + K$, $i = 1, \ldots, n$ and $X := (X_1, \ldots, X_n)$. We consider the random set— called the *Boolean model* associated with K and $C_1, \ldots, C_n$—formed by the union of K translated by the germs, that is:

$$F_n = \bigcup_{k=1}^n X_k.$$

Observe that, by construction, F_n is a random element with values in the *convex ring* $\mathcal{R}^d$ of $\mathbb{R}^d$, defined as the collection of all finite unions of convex compact subsets of $\mathbb{R}^d$.

As anticipated, our aim is to study the fluctuations of the scaled joint law of the intrinsic volumes of F_n, that is, the law of $f(X_1, \ldots, X_n)$, for $f = (f_0, \ldots, f_d)$ defined by

$$f_i(X_1, \ldots, X_n) = \frac{1}{\sqrt{n}}\big(V_i(F_n) - \mathbb{E}V_i(F_n)\big)$$

with V_i denoting the ith *intrinsic volume* (observe that the definition of f changes with n, and that such a dependence is omitted for notational clarity). We recall (see [44], Chapter 14.2, for more details) that each V_i is an additive real-valued mapping on $\mathcal{R}^d$, and that $V_d(B)$, $V_{d-1}(B)$ and $V_0(B)$ coincide, respectively, with the volume, half the surface area and the Euler characteristic of $B \in \mathcal{R}^d$. We will also use the so-called *Wills functional* $\overline{V}$, which is defined for all compact convex sets A as

$$\overline{V}(A) = \sum_{l=0}^{d} \kappa_{d-l} V_l(A),$$

with κ_{d-l} denoting the volume of the $(d - l)$-dimensional unit ball. From now on, we write

$$W := f(X)$$

and let Σ_n denote the covariance matrix of W. Finally, we let $\Sigma \in \mathbb{R}^{(d+1)\times(d+1)}$ be given by

$$(3.1) \quad \Sigma_{i,j} := \sum_{k=2}^{\infty} \frac{1}{k!} \sum_{s=i}^{d} \sum_{r=j}^{d} \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} P_{i,s}(d) P_{j,r}(d) V_s\big(K \cap (K + x_2) \cap \cdots \cap (K + x_k)\big)$$

$$\cdot V_r\big(K \cap (K + x_2) \cap \cdots \cap (K + x_k)\big) \, dx_2 \cdots dx_k, \quad i, j \in \{0, \ldots, d\},$$

where $P_{d,d}(d) := e^{-V_d(K)}$ and, for $i = 0, \ldots, d - 1$ and $s = i, \ldots, d$,

$$P_{i,s}(d) := e^{-V_d(K)}\left[\mathbb{1}_{[s=i]} + \frac{s!\kappa_s}{i!\kappa_i} \sum_{t=1}^{s-i} \frac{(-1)^t}{t!} \sum_{\substack{i \leq r_1, \ldots, r_t \leq d-1 \\ r_1 + \cdots + r_t = td + i - s}} \prod_{m=1}^{t} \frac{r_m! \kappa_{r_m}}{d! \kappa_d} V_{r_m}(K)\right].$$

The following result provides a useful upper bound on the speed of convergence of Σ_n to Σ.

THEOREM 3.1. *The matrix Σ defined above is positive definite and for all $n \in \mathbb{N}$, such that $n > \overline{V}(K)e$, and all $i, j = 0, \ldots, d$,*

$$\big|(\Sigma_n)_{i,j} - \Sigma_{i,j}\big| \leq 116 \cdot 108^d \cdot (R+1)^{4d} e^{6 \cdot 9^d \cdot (R+1)^{2d}} (d!)^2 n^{-1/d}.$$

The quantitative CLTs featured in the next statement are the main achievement of the present section. For $F \in C^2(\mathbb{R}^{d+1})$, we adopt the notation

$$M(F) := \max\{M_1(F), M_2(F)\};$$

see (1.1).

THEOREM 3.2. *Let N_Σ denote a centred $(d + 1)$-dimensional normal random vector with covariance Σ given by (3.1), and let N_{Σ_n} denote a centred normal random variable with covariance Σ_n. Let $F \in C^2(\mathbb{R}^{d+1})$. For $n > 3^{d+1}(R + 1)^d$, we have that, for $a = a(d) := \min\{1, d/2\}$,*

$$\big|\mathbb{E}F(W) - \mathbb{E}F(N_\Sigma)\big| \leq 14 \cdot M(F) \cdot 81^d \cdot (R+1)^{4d}(d+2)^3$$

$$\cdot (2^{2^d + 2d} d^{d/2})^3 (d!)^3 e^{6 \cdot 9^d (R+1)^{2d}} \big\|\Sigma^{-1}\big\|_{\mathrm{op}} n^{-a/d};$$

$$d_{\mathrm{convex}}(W, N_\Sigma) \leq 5 \cdot 10^5 \cdot 108^d \cdot (R+1)^{4d}(d+1)^7 (2^{2^d + 2d} d^{d/2})^3$$

$$\cdot e^{102 \cdot 9^d (R+1)^{2d}} (d!)^3 \max\{1, \big\|\Sigma^{-1}\big\|_{\mathrm{op}}^2\} n^{-a/d}.$$

Moreover, there exists $N \in \mathbb{N}$, *such that for all* $n \geq N$,

$$\left| \mathbb{E} F(W) - \mathbb{E} F(N_{\Sigma_n}) \right| \leq 14 \cdot M(F) \cdot 81^d \cdot (R+1)^{4d} (d+2)^3$$
$$\cdot (2^{2^d + 2d} d^{d/2})^3 (d!)^3 e^{8(2R+1)^{2d}} \left\| \Sigma_n^{-1} \right\|_{\mathrm{op}} n^{-1/2};$$

$$d_{\mathrm{convex}}(W, N_{\Sigma_n}) \leq 4 \cdot 10^5 \cdot 81^d \cdot (R+1)^{4d} (d+1)^7 (2^{2^d + 2d} d^{d/2})^3$$
$$\cdot e^{456(2R+1)^d} (d!)^3 \max\{1, \left\| \Sigma_n^{-1} \right\|_{\mathrm{op}}^2 \} n^{-1/2}.$$

REMARK 3.3. The quantity $\|\Sigma^{-1}\|_{\mathrm{op}}$ is equal to the inverse of the smallest eigenvalue of Σ, which may be bounded using techniques presented, for instance, in [30].

REMARK 3.4 (Literature review). One-dimensional quantitative CLTs in the Kolmogorov distance for the volume and the number of isolated balls of the binomial Boolean model are proved in [19] and [24], Section 6.1, respectively, in the case of a deterministic radius R (as in our setting) and in the case of i.i.d. random radii satisfying some adequate integrability assumptions. Similar bounds in the 1-Wasserstein distance for the fixed radius case are proved in [6], Section 3.2. These results are quantitative versions of classical CLTs proved for example, in [31, 36, 37]. The results from [6, 19, 24] are all obtained by using Stein's method, and yield presumably optimal rates of convergence (i.e., rates proportional to $n^{-1/2}$, as in the second part of our Theorem 3.2). The approach towards Stein's method adopted in [19] is based on the use of size-biased couplings, whereas references [6, 24] exploit discrete integration by parts formulae based on the use of the same discrete operators Δ_j as those appearing in the present work. It is in principle possible to extend our results in order to accommodate the case of random radii; for the sake of conciseness (in particular, in view of the technical nature of the proofs) we will tackle this point elsewhere. Plainly, our Theorem 3.2 also yields one-dimensional CLTs for the intrinsic volumes $V_i(F_n)$, for all $i = 0, 1, \ldots, d$; to the best of our knowledge, the CLTs in the case $0 \leq i < d$ are new even in their qualitative versions.

Multidimensional quantitative CLTs for the intrinsic volumes (and more general geometric functionals) of the Poisson-based Boolean model are proved in [22], Theorem 9.1, and [45], Section 4.2, respectively, in the case of smooth test functions, and of the convex distance adopted in our work. In both instances, the authors are able to deal with random radii satisfying adequate moment conditions, and obtain rates of convergence of the same order as those displayed in Theorem 3.2 above (to see this, observe that the inradius of the cube E_n is commensurate to $n^{1/d}$). The prefactors appearing in the bounds from [45] also display super-exponential dependence on the dimensional parameter d. The limiting covariance structure in formula (3.1) above—expressed as an infinite sum—is also featured in [22], formula (3.4), with the only difference that the order of summation there starts from $k = 1$. We stress once again that the recursive approach for dealing with the distance d_{convex} initiated in [45] is one of the most crucial ingredients of our approach.

4. An application to structures with local dependence. In this section we extend the results of [6], Section 2.3—about the normal approximation of random quantities based on structures with *local dependence*—to the multivariate setting, with specific focus on approximations in the convex distance. Proofs are gathered together in sections C and D of the *supplement* [23].

We adopt the setup of Section 1.3, fix $d \geq 1$ and denote by G a *graphical rule*, that is, a map which associates with every $x \in \mathcal{X}^n$ an undirected graph $G(x)$ on $[n] := \{1, \ldots, n\}$. We

call G *symmetric* if, for any permutation π of $[n]$ and any $(x_1, \ldots, x_n) \in \mathcal{X}^n$, the set of edges in $G(x_{\pi(1)}, \ldots, x_{\pi(n)})$ is given by

$$\{\{\pi(i), \pi(j)\} : \{i, j\} \in G(x_1, \ldots, x_n)\}.$$

For $m > n$, we say that a vector $x \in \mathcal{X}^n$ is *embedded* in $y \in \mathcal{X}^m$ if there exist distinct $i_1, \ldots, i_n \in [m]$ with $x_k = y_{i_k}$ for $1 \leq k \leq n$. We call a graphical rule G' an *extension* of G if, for any $x \in \mathcal{X}^n$ embedded in $y \in \mathcal{X}^m$, the graph $G(x)$ on $[n]$ is the naturally induced subgraph of the graph $G'(y)$ on $[m]$.

Now, taking $x, x' \in \mathcal{X}^n$ and $i \in [n]$, we let x^i be the vector obtained by replacing x_i with x_i' in the vector x. For any two distinct elements i and j of $[n]$, we let x^{ij} be the vector resulting from replacing x_i with x_i' and x_j with x_j'. We also say that the coordinates i and j are *noninteracting* under the triple (f, x, x') if

$$f(x) - f(x^j) = f(x^i) - f(x^{ij}).$$

Moreover, a graphical rule G is called an *interaction rule* for a function f if, for any $x, x' \in \mathcal{X}^n$ and $i, j \in [n]$, the event that $\{i, j\}$ is not an edge in the graphs $G(x), G(x^i), G(x^j)$ and $G(x^{ij})$ implies that i and j are noninteracting vertices under (f, x, x'). The following result, which is one of the main achievements of the present section, is a multivariate analogue of [6], Theorem 2.5, in the convex distance.

THEOREM 4.1. *Let $f : \mathcal{X}^n \to \mathbb{R}^d$ be a measurable map that admits a symmetric interaction rule G and let $M = \max_j \|\Delta_j f(X)\|$. Let G' be an arbitrary symmetric extension of G on $\mathcal{X}^{n+4}$ and set*

$$\delta := 1 + \text{degree of vertex 1 in } G'(X_1, \ldots, X_{n+4}).$$

Write $W = f(X)$ and assume $\mathbb{E}W = 0$ and $\mathbb{E}\|W\|^2 < \infty$. Let Σ denote the covariance matrix of W and N_Σ be a centred Gaussian d-dimensional vector with covariance Σ. For any $F \in C^3(\mathbb{R}^d)$, one has that

$$\left|\mathbb{E}F(W) - \mathbb{E}F(N_\Sigma)\right| \leq C\tilde{M}_2(F)\left(\mathbb{E}(M^8)\right)^{1/4}\left(\mathbb{E}(\delta^4)\right)^{1/4}n^{1/2}$$

(4.1)

$$+ \frac{M_3(F)}{12} \sum_{j=1}^n \mathbb{E}\|\Delta_j f(X)\|^3,$$

for a universal constant $C > 0$, not depending on d or n. If, in addition, Σ is positive definite, then, for any $F \in C^2(\mathbb{R}^d)$,

$$\left|\mathbb{E}F(W) - \mathbb{E}F(N_\Sigma)\right| \leq CM_1(F)\|\Sigma^{-1}\|_{\text{op}}\left(\mathbb{E}(M^8)\right)^{1/4}\left(\mathbb{E}(\delta^4)\right)^{1/4}n^{1/2}$$

(4.2)

$$+ \frac{M_2(F)\sqrt{2\pi}}{16}\|\Sigma^{-1}\|_{\text{op}} \sum_{j=1}^n \mathbb{E}\|\Delta_j f(X)\|^3,$$

and

$$d_{\text{convex}}(W, N_\Sigma) \leq Cd^4 \max(1, \|\Sigma^{-1}\|_{\text{op}}^2)$$

(4.3)

$$\cdot \max\{\left(\mathbb{E}(M^8)\right)^{1/4}\left(\mathbb{E}(\delta^4)\right)^{1/4}n^{1/2}, \gamma_1, \gamma_2, \tilde{\gamma}_3, \tilde{\gamma}_4\},$$

where $C > 0$ is a universal constant not depending on d or n, the quantities γ_1, γ_2 are defined as in Theorem 2.2 and

$$\tilde{\gamma}_3 = \left(\mathbb{E}(M^{10})\right)^{1/6}\left(\mathbb{E}(\delta^4)\right)^{1/6}n^{1/3}; \qquad \tilde{\gamma}_4 = \left(\mathbb{E}(M^{12})\right)^{1/8}\left(\mathbb{E}(\delta^4)\right)^{1/8}n^{1/4}.$$

We now present a result directly applying Theorem 4.1 to the normal approximation of nearest-neighbour statistics (cf. [6], Theorem 3.4). Such a result represents a quantitative multivariate extension of Bickel and Breiman's famous CLT for nearest neighbour statistics (see [4], as well as [6], Section 3.4). To the best of our knowledge, the content of Theorem 4.2 represents the first quantitative multivariate extension of the findings of [4].

THEOREM 4.2. *Fix integers $m, k \geq 1$. Suppose $X_1, \ldots, X_n$ are i.i.d. $\mathbb{R}^m$-valued random variables and that $\|X_1 - X_2\|$ is a continuous random variable. Let $f : (\mathbb{R}^m)^n \to \mathbb{R}^d$ be a function taking the form*

$$(4.4) \qquad f(x_1, \ldots, x_n) = \frac{1}{\sqrt{n}} \sum_{l=1}^{n} f_l(x_1, \ldots, x_n),$$

where, for each l, the value $f_l(x_1, \ldots, x_n)$ depends only on x_l and its k nearest neighbours. Suppose that $\eta_p := \max_l \mathbb{E}\|f_l(X_1, \ldots, X_n)\|^p$ is finite for some $p \geq 8$. Let $W = f(X_1, \ldots, X_n)$ and assume $\mathbb{E}W = 0$ and $\mathbb{E}\|W\|^2 < \infty$. Let Σ denote the covariance matrix of W and N_Σ be a centred Gaussian d-dimensional vector with covariance Σ. For any $F \in C^3(\mathbb{R}^d)$ with bounded second and third derivative,

$$(4.5) \qquad \left|\mathbb{E}F(W) - \mathbb{E}F(N_\Sigma)\right| \leq C \frac{\alpha(d)^3 k^4 \eta_p^{2/p}}{n^{(p-8)/(2p)}} + C \frac{\alpha(d)^3 k^3 \eta_p^{3/p}}{n^{(p-6)/(2p)}},$$

where $\alpha(d)$ is the minimum number of $60°$ cones at the origin required to cover $\mathbb{R}^d$ and C is a universal constant. The same conclusion holds if Σ is positive definite and $F \in C^2(\mathbb{R}^d)$ with bounded first and second derivative.

Furthermore, if, for some $p \geq 12$, $\eta_p := \max_l \mathbb{E}\|f_l(X_1, \ldots, X_n)\|^p$ is finite and Σ is positive definite then

$$(4.6) \qquad d_{\mathrm{convex}}(W, N_\Sigma) \leq \tilde{C} d^4 \max\left(1, \|\Sigma^{-1}\|_{\mathrm{op}}^2\right) \max\left(\frac{\alpha(d)^3 k^4 \eta_p^{2/p}}{n^{(p-8)/(2p)}}, \frac{\alpha(d)^3 k^3 \eta_p^{3/p}}{n^{(p-6)/(2p)}},\right.$$
$$\left.\frac{\alpha(d)^2 k^2 \eta_p^{2/p}}{n^{(p-4)/(2p)}}, \frac{\alpha(d)^{10/3} k^{4/3} \eta_p^{5/(3p)}}{n^{(3p-20)/(6p)}}, \frac{\alpha(d)^2 k^{5/2} \eta_p^{3/(2p)}}{n^{(p-6)/(2p)}}\right),$$

for a universal constant $\tilde{C}$, not depending on n or d.

REMARK 4.3. Our bounds for smooth test functions in (4.1) and (4.2) are of exactly the same order as the bound in [6], Theorem 2.5. The convex-distance bound (4.3) features additional terms which are somewhat difficult to compare to [6], Theorem 2.5. It is, however, possible to draw such useful comparisons in the context of the specific application covered by Theorem 4.2. The smooth distance bound in (4.5) is again of the same order as the bound in [6], Theorem 3.4. Indeed, concentrating only on the dependence on n, the bound is of order

$$(4.7) \qquad \max(\eta_p^{3/p} n^{3/p-1/2}, \eta_p^{2/p} n^{4/p-1/2}).$$

In order to extend our result in Theorem 4.2 to the convex distance case, we require a stronger integrability assumption (we need η_p to exist for some $p \geq 12$, while in our and [6] smooth distance bounds, it is enough to have it finite for some $p \geq 8$). If we assume that η_p is nondecreasing with n then our convex-distance bound (4.6) is of the same order in n as the smooth-distance bound (4.5) and the smooth-distance bound of [6], Theorem 3.4. If η_p decreases with n then we get a bound of order $\max(\eta_p^{2/p} n^{4/p-1/2}, \eta_p^{3/(2p)} n^{3/p-1/2}, \eta_p^{5/(3p)} n^{10/(3p)-1/2})$ and it might be different from (4.7) (depending on how fast η_p decreases).

Following the lines of [6], Section 3.4, one could use Theorem 4.2 in order to deduce quantitative multivariate CLTs for several classes of statistics—like for example, vertex degrees in random geometric graphs [36], Chapter 4, or dimensional estimators [27]. Details are omitted so as to keep the length of the paper within bounds.

5. Proofs of abstract results.

5.1. *Proof of Theorem* 2.1.

5.1.1. *Introduction.* Fix a function $F \in C^2(\mathbb{R}^d)$ and an $\epsilon \in \mathbb{R}$. Let $\varphi_{\epsilon^2 I}$ be the density of a d-dimensional Gaussian vector with covariance matrix $\epsilon^2 I$. Let $F_\epsilon = F \star \varphi_{\epsilon^2 I}$ and note that $F_\epsilon \in C^\infty(\mathbb{R}^d)$ and $\sup_{x \in \mathbb{R}^d} |F_\epsilon(x) - F(x)| \xrightarrow{\epsilon \to 0} 0$. In order to upper-bound $|\mathbb{E}F(W) - \mathbb{E}F(N_\Sigma)|$, we will first upper-bound $|\mathbb{E}F_\epsilon(W) - \mathbb{E}F_\epsilon(N_\Sigma)|$. This will be achieved by an application of [29], Lemma 1, which implies that there exists a function $\rho_\epsilon \in C^\infty(\mathbb{R}^d)$ solving the corresponding Stein equation for N_Σ and test function F_ϵ, that is, satisfying

$$(5.1) \qquad F_\epsilon(y) - \mathbb{E}F_\epsilon(N_\Sigma) = \langle \nabla \rho_\epsilon(y), y \rangle - \langle \mathrm{Hess}(\rho_\epsilon(y)), \Sigma \rangle_{\mathrm{H.S.}} \quad \text{for all } y \in \mathbb{R}^d.$$

We shall also apply [29], Lemma 2, stating that, for any nonnegative definite Σ,

$$(5.2) \qquad M_k(\rho_\epsilon) \le \frac{1}{k} M_k(F_\epsilon) \quad \forall k \ge 1 \quad \text{and} \quad \tilde{M}_2(\rho_\epsilon) \le \frac{1}{2} \tilde{M}_2(F_\epsilon).$$

Moreover, it follows from [8], Lemma 2.2, and [17], Proposition 2.1, that for positive-definite Σ,

$$(5.3) \qquad \tilde{M}_2(\rho_\epsilon) \le M_1(F_\epsilon) \big\| \Sigma^{-1/2} \big\|_{\mathrm{op}}; \qquad M_3(\rho_\epsilon) \le M_2(F_\epsilon) \big\| \Sigma^{-1/2} \big\|_{\mathrm{op}} \frac{\sqrt{2\pi}}{4}.$$

We now prove a lemma which will let us conclude the argument.

5.1.2. *Auxiliary lemma.*

LEMMA 5.1 (cf. Lemma 2.4 and Lemma 4.1 of [6]). *For any $\rho \in C^3(\mathbb{R}^d)$, with bounded second and third derivatives, we have*

$$\big| \mathbb{E} \langle \nabla \rho(W), W \rangle - \mathbb{E} \langle \mathrm{Hess}(\rho(W)), \Sigma \rangle_{\mathrm{H.S.}} \big|$$
$$\le \tilde{M}_2(\rho) \mathbb{E} \big\| \mathbb{E}[T | W] - \Sigma \big\|_{\mathrm{H.S.}} + \frac{M_3(\rho)}{4} \sum_{j=1}^n \mathbb{E} \big\| \Delta_j f(X) \big\|^3.$$

PROOF. For each $A \subset [n]$ and $j \notin A$, let

$$R_{A,j} := \sum_{i=1}^d \Delta_j \bigg(\frac{\partial \rho}{\partial x_i} \circ f \bigg)(X) \cdot \Delta_j f_i(X^A) \in \mathbb{R};$$

$$\tilde{R}_{A,j} = \sum_{i,k=1}^d \mathrm{Hess}_{i,k}(\rho)(W) \cdot \Delta_j f_k(X) \cdot \Delta_j f_i(X^A) \in \mathbb{R}.$$

Applying Lemma 1.1 to $g = \frac{\partial \rho}{\partial x_i} \circ f$ and $h = f_i$ and recalling that $\mathbb{E}W = 0$, we obtain:

$$(5.4) \qquad \mathbb{E} \langle \nabla \rho(W), W \rangle = \frac{1}{2} \sum_{A \subsetneq [n]} \frac{1}{\binom{n}{|A|}(n - |A|)} \sum_{j \notin A} \mathbb{E}[R_{A,j}].$$

Now, from the definition of T, we have

$$(5.5) \qquad \langle \mathrm{Hess}(\rho(W)), T \rangle_{\mathrm{H.S.}} = \frac{1}{2} \sum_{A \subsetneq [n]} \frac{1}{\binom{n}{|A|}(n - |A|)} \sum_{j \notin A} \tilde{R}_{A,j}.$$

Moreover, using Taylor's theorem and Hölder's inequality,

$$
\mathbb{E}|R_{A,j} - \tilde{R}_{A,j}| = \mathbb{E}|D\rho(f(X))[\Delta_j f(X^A)] - D\rho(f(X^j))[\Delta_j f(X^A)]
$$
$$
- D^2\rho(f(X))[\Delta_j f(X^A), \Delta_j f(X)]|
$$

(5.6)
$$
\leq \frac{M_3(\rho)}{2} \mathbb{E}[\|\Delta_j f(X)\|^2 \|\Delta_j f(X^A)\|]
$$
$$
\leq \frac{M_3(\rho)}{2} \mathbb{E}\|\Delta_j f(X)\|^3.
$$

Combining (5.4), (5.5) and (5.6), we get

$$
|\mathbb{E}\langle \nabla\rho(W), W\rangle - \mathbb{E}\langle \text{Hess}(\rho)(W), T\rangle_{\text{H.S.}}|
$$

(5.7)
$$
\leq \frac{M_3(\rho)}{4} \sum_{A \subsetneq [n]} \frac{1}{\binom{n}{|A|}(n-|A|)} \sum_{j \notin A} \mathbb{E}\|\Delta_j f(X)\|^3
$$
$$
= \frac{M_3(\rho)}{4} \sum_{j=1}^{n} \mathbb{E}\|\Delta_j f(X)\|^3.
$$

Moreover,

(5.8)
$$
|\mathbb{E}\langle \text{Hess}(\rho(W)), T - \Sigma\rangle_{\text{H.S.}}| = |\mathbb{E}\langle \text{Hess}(\rho(W)), \mathbb{E}[T - \Sigma|W]\rangle_{\text{H.S.}}|
$$
$$
\leq \tilde{M}_2(\rho)\mathbb{E}\|\mathbb{E}[T|W] - \Sigma\|_{\text{H.S.}}
$$

and the result follows by (5.7) and (5.8). □

5.1.3. *Concluding argument in the proof of Theorem* 2.1. We use the well-known formula for the derivative of a convolution and Young's convolution inequality in a manner similar to that of the proof of [29], Theorem 3. Since $\|\varphi_{\epsilon^2 I}\|_{L^1} = 1$, we obtain that $M_k(F_\epsilon) \leq M_k(F)$, for all $k \geq 1$, and $\tilde{M}_2(F_\epsilon) \leq \tilde{M}_2(F)$. The bounds (2.1), (2.2) now follow by using (5.1)–(5.3) and Lemma 5.1 and taking $\epsilon \to 0$.

5.2. *Proof of Theorem* 2.2.

5.2.1. *Introduction.* Let φ_Σ denote the density of N_Σ. Let K be a measurable convex subset of $\mathbb{R}^d$ and let $h = \mathbb{1}_K$. Fix $t \in (0, 1)$. We introduce the following smoothed version of h:

$$
h_{t,\Sigma}(y) := \int_{\mathbb{R}^d} h(\sqrt{t}z + \sqrt{1-t}\,y)\varphi_\Sigma(z)\,dz = \mathbb{E}h(\sqrt{t}N_\Sigma + \sqrt{1-t}\,y), \quad y \in \mathbb{R}^d.
$$

It follows from [45], Lemma 2.2, that

(5.9)
$$
d_{\text{convex}}(W, N_\Sigma) \leq \frac{4}{3} \sup_{h \in \mathcal{I}_d} |\mathbb{E}h_{t,\Sigma}(W) - \mathbb{E}h_{t,\Sigma}(N_\Sigma)| + \frac{20}{\sqrt{2}} d \frac{\sqrt{t}}{1-t}.
$$

In order to upper-bound (5.9), we will study the following function $f_{t,h,\Sigma} : \mathbb{R}^d \to \mathbb{R}$:

(5.10)
$$
f_{t,h,\Sigma}(y) := \frac{1}{2} \int_t^1 \frac{1}{1-s} \int_{\mathbb{R}^d} (h(\sqrt{s}z + \sqrt{1-s}\,y) - h(z))\varphi_\Sigma(z)\,dz\,ds,
$$

which, as noted in [45], page 12, and proved in [29], Lemma 1, and [34], Lemma 3.3, satisfies:

(5.11)
$$
h_{t,\Sigma}(y) - \mathbb{E}h_{t,\Sigma}(N_\Sigma) = \langle \nabla f_{t,h,\Sigma}(y), y\rangle - \langle \text{Hess}(f_{t,h,\Sigma})(y), \Sigma\rangle_{\text{H.S.}},
$$

for all $y \in \mathbb{R}^d$, i.e. solves the corresponding Stein equation for N_Σ.

Changing slightly the argument in the proof of Lemma 5.1, we note that

$$\left|\mathbb{E}[\langle \mathrm{Hess}(f_{t,h,\Sigma})(W), T - \Sigma\rangle_{\mathrm{H.S.}}]\right| = \left|\mathbb{E}[\langle \mathrm{Hess}(f_{t,h,\Sigma})(W), \mathbb{E}[T - \Sigma|W]\rangle_{\mathrm{H.S.}}]\right|$$

$$\leq \sqrt{\mathbb{E}\|\mathrm{Hess}(f_{t,h,\Sigma})(W)\|_{\mathrm{H.S.}}^2} \sqrt{\mathbb{E}\|\mathbb{E}[T - \Sigma|W]\|_{\mathrm{H.S.}}^2}.$$

By [45], Proposition 2.3,

$$\sup_{h \in \mathcal{I}_d} \mathbb{E}\|\mathrm{Hess}(f_{t,h,\Sigma})(W)\|_{\mathrm{H.S.}}^2 \leq \|\Sigma^{-1}\|_{\mathrm{op}}^2 (d^2(\log t)^2 d_{\mathrm{convex}}(W, N_\Sigma) + 530 d^{17/6}),$$

for all $t \in (0, 1)$ and therefore:

$$(5.12) \quad \begin{aligned} &\left|\mathbb{E}\langle \mathrm{Hess}(f_{t,h,\Sigma})(W), T - \Sigma\rangle_{\mathrm{H.S.}}\right| \\ &\leq \|\Sigma^{-1}\|_{\mathrm{op}} d\sqrt{(\log t)^2 d_{\mathrm{convex}}(W, N_\Sigma) + 530 d^{5/6}} \sqrt{\mathbb{E}\|\mathbb{E}[T - \Sigma|W]\|_{\mathrm{H.S.}}^2}. \end{aligned}$$

An upper bound on (5.9) will be obtained by using (5.11) and combining (5.12) with an upper-bound on $|\mathbb{E}\langle \nabla f_{t,h,\Sigma}(W), W\rangle - \langle \mathrm{Hess}(f_{t,h,\Sigma})(W), T\rangle_{\mathrm{H.S.}}|$, which we will work out in the subsequent steps of the proof. To this end, we shall follow an argument similar to that of the proof of [45], Theorem 1.2.

5.2.2. *Step* 1. As before, in the proof of Lemma 5.1, for any $A \subset [n]$ such that $j \notin A$, let

$$R_{A,j} := \sum_{i=1}^d \Delta_j \left(\frac{\partial f_{t,h,\Sigma}}{\partial x_i} \circ f\right)(X) \cdot \Delta_j f_i(X^A) \in \mathbb{R};$$

$$\tilde{R}_{A,j} := \sum_{i,k=1}^d \mathrm{Hess}_{i,k}(f_{t,h,\Sigma})(W) \cdot \Delta_j f_k(X) \cdot \Delta_j f_i(X^A) \in \mathbb{R}.$$

For $A \subsetneq [n]$, let $k_{n,A} = \frac{1}{\binom{n}{|A|}(n-|A|)}$. It follows that

$$(5.13) \quad \left|\mathbb{E}\langle \nabla f_{t,h,\Sigma}(W), W\rangle - \langle \mathrm{Hess}(f_{t,h,\Sigma}), T\rangle_{\mathrm{H.S.}}\right| = \frac{1}{2}\left|\sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \mathbb{E}[R_{A,j} - \tilde{R}_{A,j}]\right|.$$

Now, using repeatedly Taylor's theorem with integral remainder (in identities (∗) below), we obtain, for $j \notin A$, that

$$\begin{aligned} &\mathbb{E}[\tilde{R}_{A,j} - R_{A,j}] \\ &= \mathbb{E}[(Df_{t,h,\Sigma}(f(X^j)) - Df_{t,h,\Sigma}(f(X)))[\Delta_j f(X^A)] \\ &\quad - D^2 f_{t,h,\Sigma}(f(X))[\Delta_j f(X^A), -\Delta_j f(X)]] \\ &\stackrel{(*)}{=} \mathbb{E}\int_0^1 \{D^2 f_{t,h,\Sigma}(f(X) - u\Delta_j f(X))[\Delta_j f(X^A), -\Delta_j f(X)] \\ &\quad - D^2 f_{t,h,\Sigma}(f(X))[\Delta_j f(X^A), -\Delta_j f(X)]\} \, du \\ &= \mathbb{E}\int_0^1 (D^2 f_{t,h,\Sigma}(f(X) - u\Delta_j f(X)) - D^2 f_{t,h,\Sigma}(f(X))) \\ &\quad [\Delta_j f(X^A), -\Delta_j f(X)] \, du \\ &\stackrel{(*)}{=} \mathbb{E}\int_0^1 \int_0^1 D^3 f_{t,h,\Sigma}(f(X) - uv\Delta_j f(X)) \\ &\quad [\Delta_j f(X^A), \Delta_j f(X), u\Delta_j f(X)] \, dv \, du \end{aligned}$$

$$(5.14) \qquad = \mathbb{E} \int_0^1 \int_0^1 D^3 f_{t,h,\Sigma}(f(X) - v\Delta_j f(X))$$

$$[\Delta_j f(X^A), \Delta_j f(X), u\Delta_j f(X)] \, dv \, du$$

$$+ \mathbb{E} \int_0^1 \int_0^1 (D^3 f_{t,h,\Sigma}(f(X) - uv\Delta_j f(X)) - D^3 f_{t,h,\Sigma}(f(X) - v\Delta_j f(X)))$$

$$[\Delta_j f(X^A), \Delta_j f(X), u\Delta_j f(X)] \, dv \, du$$

$$\overset{(*)}{=} \frac{1}{2} \mathbb{E}(D^2 f_{t,h,\Sigma}(f(X)) - D^2 f_{t,h,\Sigma}(f(X^j)))[\Delta_j f(X^A), \Delta_j f(X)]$$

$$+ \mathbb{E} \int_0^1 \int_0^1 (D^3 f_{t,h,\Sigma}(f(X) - uv\Delta_j f(X)) - D^3 f_{t,h,\Sigma}(f(X) - v\Delta_j f(X)))$$

$$[\Delta_j f(X^A), \Delta_j f(X), u\Delta_j f(X)] \, dv \, du$$

$$= \mathbb{E} \int_0^1 \int_0^1 (D^3 f_{t,h,\Sigma}(f(X) - uv\Delta_j f(X)) - D^3 f_{t,h,\Sigma}(f(X) - v\Delta_j f(X)))$$

$$[\Delta_j f(X^A), \Delta_j f(X), u\Delta_j f(X)] \, dv \, du$$

$$=: J(j, A).$$

The second to last identity holds because, for $j \notin A$,

$$(X, X^j, X^A, X^{A\cup\{j\}}) \overset{\mathcal{D}}{=} (X^j, X, X^{A\cup\{j\}}, X^A)$$

and so

$$\mathbb{E}(D^2 f_{t,h,\Sigma}(f(X)) - D^2 f_{t,h,\Sigma}(f(X^j)))[\Delta_j f(X^A), \Delta_j f(X)]$$

$$= \mathbb{E}D^2 f_{t,h,\Sigma}(f(X))[f(X^A) - f(X^{A\cup\{j\}}), f(X) - f(X^j)]$$

$$- \mathbb{E}D^2 f_{t,h,\Sigma}(f(X^j))[f(X^{A\cup\{j\}}) - f(X^A), f(X^j) - f(X)]$$

$$= 0.$$

Now, using (5.10) and [45], (2.2), we obtain that, for $J(j, A)$, defined in (5.14),

$$J(j, A) = -\frac{1}{2} \sum_{k,l,m=1}^d \mathbb{E} \int_0^1 \int_0^1 \int_t^1 \frac{\sqrt{1-s}}{s^{3/2}} \int_{\mathbb{R}^d} [h(\sqrt{s}z + \sqrt{1-s}(f(X) - uv\Delta_j f(X)))$$

$$- h(\sqrt{s}z + \sqrt{1-s}(f(X) - v\Delta_j f(X)))]$$

$$\cdot \frac{\partial^3 \varphi_\Sigma}{\partial y_k \, \partial y_l \, \partial y_m}(z) u(\Delta_j f(X^A))_k (\Delta_j f(X))_l (\Delta_j f(X))_m \, dz \, ds \, dv \, du.$$

Using the abbreviation

$$U_{k,l,m}$$

$$:= \sup_{\substack{z \in \mathbb{R}^d \\ s,u \in [0,1]}} \mathbb{E}\Bigg\{ \sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \int_0^1 \Big| [h(\sqrt{s}z + \sqrt{1-s}(f(X) - uv\Delta_j f(X)))$$

$$- h(\sqrt{s}z + \sqrt{1-s}(f(X) - v\Delta_j f(X)))] (\Delta_j f(X^A))_k (\Delta_j f(X))_l (\Delta_j f(X))_m \Big| dv \Bigg\},$$

for $k, l, m \in [d]$, and the Cauchy–Schwarz inequality, we obtain

$$
(5.15) \quad
\begin{aligned}
\frac{1}{2}\Bigg| \sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} J(j, A) \Bigg| \\
\leq \frac{1}{2\sqrt{t}} \sum_{k,l,m=1}^{d} \int_{\mathbb{R}^d} \left| \frac{\partial^3 \varphi_\Sigma}{\partial y_k \, \partial y_l \, \partial y_m}(z) \right| dz \, U_{k,l,m} \\
\leq \frac{1}{2\sqrt{t}} \int_{\mathbb{R}^d} \left(\sum_{k,l,m=1}^{d} \left(\frac{\partial^3 \varphi_\Sigma}{\partial y_k \, \partial y_l \, \partial y_m}(z) \right)^2 \right)^{1/2} dz \left(\sum_{k,l,m=1}^{d} (U_{k,l,m})^2 \right)^{1/2}.
\end{aligned}
$$

By [45], page 22 and (3.9),

$$
(5.16) \quad \int_{\mathbb{R}^d} \left(\sum_{k,l,m=1}^{d} \left(\frac{\partial^3 \varphi_\Sigma}{\partial y_k \, \partial y_l \, \partial y_m}(z) \right)^2 \right)^{1/2} dz \leq \sqrt{6} d^{3/2} \| \Sigma^{-1} \|_{\mathrm{op}}^{3/2}.
$$

Therefore, by (5.15), in order to upper-bound (5.13), it remains to upper-bound $U_{k,l,m}$, for $k, l, m \in [d]$.

5.2.3. *Step 2—an upper bound on $U_{k,l,m}$.* For j such that $\Delta_j f(X) \neq 0$, we define $r(\Delta_j f(X)) := \frac{1}{\|\Delta_j f(X)\|} \Delta_j f(X)$. We also let $w = v \| \Delta_j f(X) \|$. We obtain

$$
(5.17) \quad
\begin{aligned}
U_{k,l,m} \leq \sup_{\substack{z \in \mathbb{R}^d \\ s, u \in [0,1]}} \mathbb{E}\Bigg\{ \sum_{A \subsetneq [n]} k_{n,A} \\
\cdot \sum_{j \notin A} \int_0^{\|\Delta_j f(X)\|} \Bigg| h\big(\sqrt{s}z + \sqrt{1-s}(f(X) - uwr(\Delta_j f(X)))\big) \\
- h\big(\sqrt{s}z + \sqrt{1-s}(f(X) - wr(\Delta_j f(X)))\big) \Bigg| dw \\
\cdot \mathbb{1}_{[\|\Delta_j f(X)\| \leq 1]} \Bigg| (\Delta_j f(X^A))_k \frac{(\Delta_j f(X))_l}{\|\Delta_j f(X)\|} (\Delta_j f(X))_m \Bigg| \Bigg\} \\
+ \sup_{\substack{z \in \mathbb{R}^d \\ s, u \in [0,1]}} \mathbb{E}\Bigg\{ \sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \int_0^1 \Bigg| \big[h\big(\sqrt{s}z + \sqrt{1-s}(f(X) - uv\Delta_j f(X))\big) \\
- h\big(\sqrt{s}z + \sqrt{1-s}(f(X) - v\Delta_j f(X))\big) \big] \\
\cdot \mathbb{1}_{[\|\Delta_j f(X)\| > 1]} (\Delta_j f(X^A))_k (\Delta_j f(X))_l (\Delta_j f(X))_m \Bigg| dv \Bigg\} \\
=: U_{k,l,m}^{(1)} + U_{k,l,m}^{(2)}.
\end{aligned}
$$

Recall that $h(\cdot) = \mathbb{1}_{[\cdot \in K]}$ for a measurable convex set $K \subseteq \mathbb{R}^d$. It follows that

$$
\begin{aligned}
U_{k,l,m}^{(2)} &\leq \sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \mathbb{E}\big\{ \mathbb{1}_{[\|\Delta_j f(X)\| > 1]} \big| (\Delta_j f(X^A))_k (\Delta_j f(X))_l (\Delta_j f(X))_m \big| \big\} \\
&\leq \sum_{A \subsetneq [n]} k_{n,A} \mathbb{E}\big\{ \| \Delta_j f(X) \| \big| (\Delta_j f(X^A))_k (\Delta_j f(X))_l (\Delta_j f(X))_m \big| \big\}
\end{aligned}
$$

and so

$$\sqrt{\sum_{k,l,m=1}^{d} (U_{k,l,m}^{(2)})^2}$$

$$\leq \sum_{k,l,m=1}^{d} |U_{k,l,m}^{(2)}|$$

(5.18)

$$\leq \sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \sum_{k,l,m=1}^{d} \mathbb{E}[\|\Delta_j f(X)\| |(\Delta_j f(X^A))_k (\Delta_j f(X))_l (\Delta_j f(X))_m|]$$

$$\leq d^{3/2} \sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \mathbb{E}[\|\Delta_j f(X)\|^3 \|\Delta_j f(X^A)\|]$$

$$\leq d^{3/2} \sum_{j=1}^{n} \mathbb{E}\|\Delta_j f(X)\|^4 = d^{3/2} \gamma_2^2.$$

Now, we need to bound $\sum_{k,l,m=1}^{d} (U_{k,l,m}^{(1)})^2$. We define $K_{s,z} := \frac{1}{\sqrt{1-s}}(K - \sqrt{s}z)$. Then, for $v, u \in [0, 1]$ and r and w defined at the top of this subsection,

$$\left|h(\sqrt{s}z + \sqrt{1-s}(f(X) - uwr(\Delta_j f(X)))) - h(\sqrt{s}z + \sqrt{1-s}(f(X) - wr(\Delta_j f(X))))\right|$$

$$= |\mathbb{1}_{[f(X) - uwr(\Delta_j f(X)) \in K_{s,z}]} - \mathbb{1}_{[f(X) - wr(\Delta_j f(X)) \in K_{s,z}]}|$$

$$\leq \mathbb{1}_{[\text{dist}(f(X), \partial K_{s,z}) \leq w]},$$

where $\partial K_{s,z}$ is the boundary of $K_{s,z}$ and $\text{dist}(f(X), \partial K_{s,z}) := \inf_{y \in \partial K_{s,z}} \|y - f(X)\|$. Thus, we have that

$$U_{k,l,m}^{(1)} \leq \sup_{z \in \mathbb{R}^d, s \in [0,1]} \mathbb{E}\left\{ \sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \int_0^1 \mathbb{1}_{[\text{dist}(f(X), \partial K_{s,z}) \leq w]} \mathbb{1}_{[w \leq \|\Delta_j f(X)\|]} \right.$$

$$\left. \cdot |(\Delta_j f(X^A))_l (\Delta_j f(X))_m| \, dw \right\}$$

$$\leq \sup_{\substack{z \in \mathbb{R}^d \\ s \in [0,1]}} \left\{ \sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \int_0^1 \mathbb{P}[\text{dist}(f(X), \partial K_{s,z}) \leq w] \right.$$

(5.19)

$$\left. \cdot \mathbb{E}[\mathbb{1}_{[w \leq \|\Delta_j f(X)\|]} |(\Delta_j f(X^A))_l (\Delta_j f(X))_m|] \, dw \right\}$$

$$+ \sup_{z \in \mathbb{R}^d, s \in [0,1]} \int_0^1 \mathbb{E}\left\{ \mathbb{1}_{[\text{dist}(f(X), \partial K_{s,z}) \leq w]} \sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \left(\mathbb{1}_{[w \leq \|\Delta_j f(X)\|]} \right. \right.$$

$$\left. \cdot |(\Delta_j f(X^A))_l (\Delta_j f(X))_m| \right.$$

$$\left. \left. - \mathbb{E}[\mathbb{1}_{[w \leq \|\Delta_j f(X)\|]} |(\Delta_j f(X^A))_l (\Delta_j f(X))_m|] \right) \right\} dw$$

$$=: R_{l,m}^{(1)} + R_{l,m}^{(2)}.$$

5.2.4. *Step* 3—*upper bounds on* $R_{l,m}^{(1)}$ *and* $R_{l,m}^{(2)}$. Now, we apply [45], Lemma 3.3, to obtain

$$\sqrt{d \sum_{l,m=1}^{d} (R_{l,m}^{(1)})^2}$$

$$\leq \sqrt{d} \sum_{l,m=1}^{d} |R_{l,m}^{(1)}|$$

$$\leq 2d \|\Sigma^{-1/2}\|_{\mathrm{op}}$$

$$\cdot \sum_{l,m=1}^{d} \sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \mathbb{E}\left\{ \int_0^1 w \mathbb{1}_{[w \leq \|\Delta_j f(X)\|]} \, dw |(\Delta_j f(X^A))_l (\Delta_j f(X))_m| \right\}$$

$$+ 2 d_{\mathrm{convex}}(W, N_\Sigma) \sqrt{d} \sum_{l,m=1}^{d} \sum_{A \subsetneq [n]} k_{n,A}$$

$$\cdot \sum_{j \notin A} \mathbb{E}\left\{ \int_0^1 \mathbb{1}_{[w \leq \|\Delta_j f(X)\|]} \, dw |(\Delta_j f(X^A))_l (\Delta_j f(X))_m| \right\}$$

(5.20)

$$\leq d \|\Sigma^{-1/2}\|_{\mathrm{op}} \sum_{l,m=1}^{d} \sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \mathbb{E}\{\|\Delta_j f(X)\|^2 |(\Delta_j f(X^A))_l (\Delta_j f(X))_m|\}$$

$$+ 2 d_{\mathrm{convex}}(W, N_\Sigma) \sqrt{d}$$

$$\cdot \sum_{l,m=1}^{d} \sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \mathbb{E}\{\|\Delta_j f(X)\| |(\Delta_j f(X^A))_l (\Delta_j f(X))_m|\}$$

$$\leq d^2 \|\Sigma^{-1/2}\|_{\mathrm{op}} \sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \mathbb{E}[\|\Delta_j f(X^A)\| \|\Delta_j f(X)\|^3]$$

$$+ \sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} 2 d_{\mathrm{convex}}(W, N_\Sigma) d^{3/2} \mathbb{E}[\|\Delta_j f(X^A)\| \|\Delta_j f(X)\|^2]$$

$$\leq d^2 \|\Sigma^{-1/2}\|_{\mathrm{op}} \sum_{j=1}^{n} \mathbb{E}\|\Delta_j f(X)\|^4 + 2 d_{\mathrm{convex}}(W, N_\Sigma) d^{3/2} \sum_{j=1}^{n} \mathbb{E}\|\Delta_j f(X)\|^3$$

$$= d^2 \|\Sigma^{-1/2}\|_{\mathrm{op}} \gamma_2^2 + 2 d_{\mathrm{convex}}(W, N_\Sigma) d^{3/2} \gamma_1.$$

For $R_{l,m}^{(2)}$ we apply the Cauchy–Schwarz inequality and [45], Lemma 3.3, to obtain

$$R_{l,m}^{(2)} \leq \sup_{z \in \mathbb{R}^d, s \in [0,1]} \int_0^1 \mathbb{P}\big(\mathrm{dist}(f(X), \partial K_{s,z}) \leq w\big)^{1/2}$$

$$\cdot \left(\mathbb{E}\bigg| \sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \mathbb{E}\big[\mathbb{1}_{[w \leq \|\Delta_j f(X)\|]} |(\Delta_j f(X^A))_l (\Delta_j f(X))_m| \, \big| \, X \big] \right.$$

(5.21) $$\left. - \sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \mathbb{E}\{ \mathbb{1}_{[w \leq \|\Delta_j f(X)\|]} |(\Delta_j f(X^A))_l (\Delta_j f(X))_m| \} \bigg|^2 \right)^{1/2} dw$$

$$\leq \int_0^1 (2d_{\mathrm{convex}}(W, N_\Sigma) + 2\sqrt{d}\|\Sigma^{-1/2}\|_{\mathrm{op}} w)^{1/2}$$

$$\cdot \left(\mathrm{Var}\Big\{ \sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \mathbb{E}\Big[\mathbb{1}_{[w \leq \|\Delta_j f(X)\|]} |(\Delta_j f(X^A))_l (\Delta_j f(X))_m| \,\big|\, X\Big]\Big\}\right)^{1/2} dw$$

$$\leq (2d_{\mathrm{convex}}(W, N_\Sigma) V_{l,m}^{(1)} + 2\sqrt{d}\|\Sigma^{-1/2}\|_{\mathrm{op}} V_{l,m}^{(2)})^{1/2},$$

by the Cauchy–Schwarz inequality, where

$$V_{l,m}^{(1)} := \int_0^1 \mathrm{Var}\Big\{ \sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \mathbb{E}\Big[\mathbb{1}_{[w \leq \|\Delta_j f(X)\|]} |(\Delta_j f(X^A))_l (\Delta_j f(X))_m| \,\big|\, X\Big]\Big\} dw$$

(5.22)

$$V_{l,m}^{(2)} := \int_0^1 w \, \mathrm{Var}\Big\{ \sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \mathbb{E}\Big[\mathbb{1}_{[w \leq \|\Delta_j f(X)\|]} |(\Delta_j f(X^A))_l (\Delta_j f(X))_m| \,\big|\, X\Big]\Big\} dw.$$

The existence of the variances in the definition of $V_{l,m}^{(1)}$ and $V_{l,m}^{(2)}$ is proved below, in (5.28).

5.2.5. *Step* 4—*upper bounds on* $V_{l,m}^{(1)}$ *and* $V_{l,m}^{(2)}$. We first compute the difference operators. Note that, for all $i, j \in [n]$,

$$|\tilde{\Delta}_i \mathbb{1}_{[w \leq \|\Delta_j f(X)\|]}| = |\mathbb{1}_{[w \leq \|\Delta_j f(X)\| - \tilde{\Delta}_i \|\Delta_j f(X)\|]} - \mathbb{1}_{[w \leq \|\Delta_j f(X)\|]}|$$

$$\leq \mathbb{1}_{[w \leq \|\Delta_j f(X)\| + |\tilde{\Delta}_i \|\Delta_j f(X)\||]};$$

$$|\tilde{\Delta}_i \|\Delta_j f(X)\|| = |\|\Delta_j f(X) - \tilde{\Delta}_i \Delta_j f(X)\| - \|\Delta_j f(X)\|| \leq \|\tilde{\Delta}_i \Delta_j f(X)\|.$$

Hence

$$|\tilde{\Delta}_i \mathbb{1}_{[w \leq \|\Delta_j f(X)\|]}| \leq \mathbb{1}_{[w \leq \|\Delta_j f(X)\| + \|\tilde{\Delta}_i \Delta_j f(X)\|]} \mathbb{1}_{[\tilde{\Delta}_i \Delta_j f(X) \neq 0]}.$$

Therefore, for all $i \in [n]$,

$$\left(\sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \tilde{\Delta}_i \big[\mathbb{1}_{[w \leq \|\Delta_j f(X)\|]} |(\Delta_j f(X^A))_l (\Delta_j f(X))_m|\big] \right)^2$$

$$\leq \Big\{ \sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \Big[\mathbb{1}_{[w \leq \|\Delta_j f(X)\| + \|\tilde{\Delta}_i \Delta_j f(X)\|]}$$

$$\cdot \mathbb{1}_{[\tilde{\Delta}_i \Delta_j f(X) \neq 0]} \big(|(\Delta_j f(X^A))_l (\Delta_j f(X))_m|$$

$$+ |\tilde{\Delta}_i |(\Delta_j f(X^A))_l (\Delta_j f(X))_m||\big)$$

$$+ \mathbb{1}_{[w \leq \|\Delta_j f(X)\|]} |\tilde{\Delta}_i |(\Delta_j f(X^A))_l (\Delta_j f(X))_m||\Big] \Big\}^2$$

(5.23)

$$\leq 3 \sum_{A_1, A_2 \subsetneq [n]} k_{n,A_1} k_{n,A_2} \sum_{j_1 \notin A_1} \sum_{j_2 \notin A_2} \big[\mathbb{1}_{[\tilde{\Delta}_i \Delta_{j_1} f(X) \neq 0, \tilde{\Delta}_i \Delta_{j_2} f(X) \neq 0]}$$

$$\cdot \mathbb{1}_{[w \leq \min\{\|\Delta_{j_1} f(X)\| + \|\tilde{\Delta}_i \Delta_{j_1} f(X)\|, \|\Delta_{j_2} f(X)\| + \|\tilde{\Delta}_i \Delta_{j_2} f(X)\|\}]}$$

$$\cdot \big(|(\Delta_{j_1} f(X^A))_l (\Delta_{j_1} f(X))_m| \cdot |(\Delta_{j_2} f(X^A))_l (\Delta_{j_2} f(X))_m|$$

$$+ |\tilde{\Delta}_i |(\Delta_{j_1} f(X^A))_l (\Delta_{j_1} f(X))_m|| \cdot |\tilde{\Delta}_i |(\Delta_{j_2} f(X^A))_l (\Delta_{j_2} f(X))_m||\big)\big]$$

$$+ 3 \sum_{A_1, A_2 \subsetneq [n]} k_{n,A_1} k_{n,A_2} \sum_{j_1 \notin A_1} \sum_{j_2 \notin A_2} \big[\mathbb{1}_{[w \leq \min\{\|\Delta_{j_1} f(X)\|, \|\Delta_{j_2} f(X)\|\}]}$$

$$\cdot |\tilde{\Delta}_i |(\Delta_{j_1} f(X^A))_l (\Delta_{j_1} f(X))_m|| \cdot |\tilde{\Delta}_i |(\Delta_{j_2} f(X^A))_l (\Delta_{j_2} f(X))_m||.$$

We now use the fact that, for any $g : \mathcal{X}^{2n} \to \mathbb{R}$ and $U = g(X, X')$, $\mathrm{Var}(\mathbb{E}(U|X)) \le \mathbb{E}(\mathrm{Var}(U|X'))$ (see [6], Lemma 4.4). Together with the Efron–Stein inequality [48], this implies that, for such U,

$$(5.24) \quad \mathrm{Var}(\mathbb{E}(U|X)) \le \mathbb{E}(\mathrm{Var}(U|X')) \le \mathbb{E}\left\{ \frac{1}{2} \sum_{i=1}^{n} \mathbb{E}[(\tilde{\Delta}_i U)^2 \mid X'] \right\} = \frac{1}{2} \sum_{i=1}^{n} \mathbb{E}[(\tilde{\Delta}_i U)^2].$$

The general fact (5.24) and the inequality (5.23) together lead to the following estimates:

$$
\begin{aligned}
V_{l,m}^{(1)} \le{}& \frac{3}{2} \sum_{i=1}^{n} \sum_{A_1, A_2 \subsetneq [n]} k_{n,A_1} k_{n,A_2} \sum_{j_1 \notin A_1} \sum_{j_2 \notin A_2} \mathbb{E}\big\{ \mathbb{1}_{[\tilde{\Delta}_i \Delta_{j_1} f(X) \neq 0, \tilde{\Delta}_i \Delta_{j_2} f(X) \neq 0]} \\
& \cdot \min\big[\|\Delta_{j_1} f(X)\| + \|\tilde{\Delta}_i \Delta_{j_1} f(X)\|, \|\Delta_{j_2} f(X)\| + \|\tilde{\Delta}_i \Delta_{j_2} f(X)\| \big] \\
& \cdot \big(|(\Delta_{j_1} f(X^A))_l (\Delta_{j_1} f(X))_m| \cdot |(\Delta_{j_2} f(X^A))_l (\Delta_{j_2} f(X))_m| \\
& + |\tilde{\Delta}_i| (\Delta_{j_1} f(X^A))_l (\Delta_{j_1} f(X))_m|| \cdot |\tilde{\Delta}_i| (\Delta_{j_2} f(X^A))_l (\Delta_{j_2} f(X))_m||) \big\} \\
& + \frac{3}{2} \sum_{i=1}^{n} \sum_{A_1, A_2 \subsetneq [n]} k_{n,A_1} k_{n,A_2} \sum_{j_1 \notin A_1} \sum_{j_2 \notin A_2} \mathbb{E}\big\{ \min\big[\|\Delta_{j_1} f(X)\|, \|\Delta_{j_2} f(X)\| \big] \\
& \cdot |\tilde{\Delta}_i| (\Delta_{j_1} f(X^A))_l (\Delta_{j_1} f(X))_m|| \cdot |\tilde{\Delta}_i| (\Delta_{j_2} f(X^A))_l (\Delta_{j_2} f(X))_m|| \big\};
\end{aligned}
$$

$$
\begin{aligned}
V_{l,m}^{(2)} \le{}& \frac{3}{2} \sum_{i=1}^{n} \sum_{A_1, A_2 \subsetneq [n]} k_{n,A_1} k_{n,A_2} \sum_{j_1 \notin A_1} \sum_{j_2 \notin A_2} \mathbb{E}\big\{ \mathbb{1}_{[\tilde{\Delta}_i \Delta_{j_1} f(X) \neq 0, \tilde{\Delta}_i \Delta_{j_2} f(X) \neq 0]} \\
& \cdot \min\big[\|\Delta_{j_1} f(X)\|^2 + \|\tilde{\Delta}_i \Delta_{j_1} f(X)\|^2, \|\Delta_{j_2} f(X)\|^2 + \|\tilde{\Delta}_i \Delta_{j_2} f(X)\|^2 \big] \\
& \cdot \big(|(\Delta_{j_1} f(X^A))_l (\Delta_{j_1} f(X))_m| \cdot |(\Delta_{j_2} f(X^A))_l (\Delta_{j_2} f(X))_m| \\
& + |\tilde{\Delta}_i| (\Delta_{j_1} f(X^A))_l (\Delta_{j_1} f(X))_m|| \cdot |\tilde{\Delta}_i| (\Delta_{j_2} f(X^A))_l (\Delta_{j_2} f(X))_m||) \big\} \\
& + \frac{3}{4} \sum_{i=1}^{n} \sum_{A_1, A_2 \subsetneq [n]} k_{n,A_1} k_{n,A_2} \sum_{j_1 \notin A_1} \sum_{j_2 \notin A_2} \mathbb{E}\big\{ \min\big[\|\Delta_{j_1} f(X)\|^2, \|\Delta_{j_2} f(X)\|^2 \big] \\
& \cdot |\tilde{\Delta}_i| (\Delta_{j_1} f(X^A))_l (\Delta_{j_1} f(X))_m|| \cdot |\tilde{\Delta}_i| (\Delta_{j_2} f(X^A))_l (\Delta_{j_2} f(X))_m|| \big\}.
\end{aligned}
$$

This, in turn, implies that, for $p = 1, 2$

$$
\begin{aligned}
V_{l,m}^{(p)} \le{}& \frac{3}{2} \sum_{i=1}^{n} \mathbb{E}\bigg[\bigg(\sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \mathbb{1}_{[\tilde{\Delta}_i \Delta_j f(X) \neq 0]} \sqrt{\|\Delta_j f(X)\|^p + \|\tilde{\Delta}_i \Delta_j f(X)\|^p} \\
& \cdot |(\Delta_j f(X^A))_l (\Delta_j f(X))_m| \bigg)^2 \bigg] \\
& + \frac{3}{2} \sum_{i=1}^{n} \mathbb{E}\bigg[\bigg(\sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \mathbb{1}_{[\tilde{\Delta}_i \Delta_j f(X) \neq 0]} \sqrt{\|\Delta_j f(X)\|^p + \|\tilde{\Delta}_i \Delta_j f(X)\|^p} \\
& \cdot |\tilde{\Delta}_i| (\Delta_j f(X^A))_l (\Delta_j f(X))_m|| \bigg)^2 \bigg] \\
(5.25) \quad & + \frac{3}{2p} \sum_{i=1}^{n} \mathbb{E}\bigg[\bigg(\sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \sqrt{\|\Delta_j f(X)\|^p} |\tilde{\Delta}_i| (\Delta_j f(X^A))_l (\Delta_j f(X))_m|| \bigg)^2 \bigg]
\end{aligned}
$$

$$\leq \frac{3}{2} \sum_{i=1}^{n} \mathbb{E}\Bigg[\bigg(\sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \mathbb{1}_{[\tilde{\Delta}_i \Delta_j f(X) \neq 0]} \sqrt{\|\Delta_j f(X)\|^p + \|\tilde{\Delta}_i \Delta_j f(X)\|^p}$$

$$\cdot \left|(\Delta_j f(X^A))_l (\Delta_j f(X))_m|\right)^2\Bigg]$$

$$+ \left(\frac{3}{2} + \frac{3}{2p}\right) \sum_{i=1}^{n} \mathbb{E}\Bigg[\bigg(\sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \sqrt{\|\Delta_j f(X)\|^p + \|\tilde{\Delta}_i \Delta_j f(X)\|^p}$$

$$\cdot \left|\tilde{\Delta}_i |(\Delta_j f(X^A))_l (\Delta_j f(X))_m||\right)^2\Bigg].$$

Now, note that

$$(5.26) \quad \begin{aligned} &|\tilde{\Delta}_i |(\Delta_j f(X^A))_l (\Delta_j f(X))_m|| \\ &\quad \leq |\tilde{\Delta}_i ((\Delta_j f(X^A))_l (\Delta_j f(X))_m)| \\ &\quad = |\tilde{\Delta}_i (\Delta_j f(X^A))_l (\Delta_j f(X))_m + (\Delta_j f(X^A))_l \tilde{\Delta}_i (\Delta_j f(X))_m \\ &\quad\quad - \tilde{\Delta}_i (\Delta_j f(X^A))_l \tilde{\Delta}_i (\Delta_j f(X))_m| \end{aligned}$$

and therefore, using (5.25), for $p = 1, 2$,

$$V_{l,m}^{(p)} \leq \frac{3}{2} \sum_{i=1}^{n} \mathbb{E}\Bigg[\bigg(\sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \mathbb{1}_{[\tilde{\Delta}_i \Delta_j f(X) \neq 0]} \sqrt{\|\Delta_j f(X)\|^p + \|\tilde{\Delta}_i \Delta_j f(X)\|^p}$$

$$\cdot \left|(\Delta_j f(X^A))_l (\Delta_j f(X))_m|\right)^2\Bigg]$$

$$+ \left(\frac{9}{2} + \frac{9}{2p}\right) \sum_{i=1}^{n} \mathbb{E}\Bigg[\bigg(\sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \sqrt{\|\Delta_j f(X)\|^p + \|\tilde{\Delta}_i \Delta_j f(X)\|^p}$$

$$\cdot \left|\tilde{\Delta}_i (\Delta_j f(X^A))_l (\Delta_j f(X))_m|\right)^2\Bigg]$$

$$+ \left(\frac{9}{2} + \frac{9}{2p}\right) \sum_{i=1}^{n} \mathbb{E}\Bigg[\bigg(\sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \sqrt{\|\Delta_j f(X)\|^p + \|\tilde{\Delta}_i \Delta_j f(X)\|^p}$$

$$\cdot \left|(\Delta_j f(X^A))_l \tilde{\Delta}_i (\Delta_j f(X))_m|\right)^2\Bigg]$$

$$+ \left(\frac{9}{2} + \frac{9}{2p}\right) \sum_{i=1}^{n} \mathbb{E}\Bigg[\bigg(\sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \sqrt{\|\Delta_j f(X)\|^p + \|\tilde{\Delta}_i \Delta_j f(X)\|^p}$$

$$\cdot \left|\tilde{\Delta}_i (\Delta_j f(X^A))_l \tilde{\Delta}_i (\Delta_j f(X))_m|\right)^2\Bigg].$$

Now,

$$\frac{3}{2} \sum_{l,m=1}^{d} \sum_{i=1}^{n} \mathbb{E}\Bigg[\bigg(\sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \mathbb{1}_{[\tilde{\Delta}_i \Delta_j f(X) \neq 0]} \sqrt{\|\Delta_j f(X)\|^p + \|\tilde{\Delta}_i \Delta_j f(X)\|^p}$$

$$\cdot \left|(\Delta_j f(X^A))_l (\Delta_j f(X))_m|\right)^2\Bigg]$$

$$\leq \frac{3}{2}\sum_{i=1}^{n}\mathbb{E}\Bigg[\Bigg(\sum_{A\subsetneq[n]}k_{n,A}\sum_{j\notin A}\mathbb{1}_{[\tilde{\Delta}_i\Delta_j f(X)\neq 0]}\sqrt{\|\Delta_j f(X)\|^p + \|\tilde{\Delta}_i\Delta_j f(X)\|^p}$$

$$\cdot \sum_{l,m=1}^{d}|(\Delta_j f(X^A))_l(\Delta_j f(X))_m|\Bigg)^2\Bigg]$$

$$\leq \frac{3}{2}d^2\sum_{i=1}^{n}\mathbb{E}\Bigg[\Bigg(\sum_{A\subsetneq[n]}k_{n,A}\sum_{j\notin A}\mathbb{1}_{[\tilde{\Delta}_i\Delta_j f(X)\neq 0]}\sqrt{\|\Delta_j f(X)\|^p + \|\tilde{\Delta}_i\Delta_j f(X)\|^p}$$

$$\cdot \|\Delta_j f(X^A)\|\|\Delta_j f(X)\|\Bigg)^2\Bigg]$$

and similar inequalities hold for the other terms. Therefore, for $p = 1, 2$,

$$\sum_{l,m=1}^{d}V_{l,m}^{(p)}$$

$$\leq \frac{3}{2}d^2\sum_{i=1}^{n}\mathbb{E}\Bigg[\Bigg(\sum_{A\subsetneq[n]}k_{n,A}\sum_{j\notin A}\mathbb{1}[\tilde{\Delta}_i\Delta_j f(X)\neq 0]\sqrt{\|\Delta_j f(X)\|^p + \|\tilde{\Delta}_i\Delta_j f(X)\|^p}$$

$$\cdot \|\Delta_j f(X^A)\|\|\Delta_j f(X)\|\Bigg)^2\Bigg]$$

$$+ d^2\Big(\frac{9}{2} + \frac{9}{2p}\Big)\sum_{i=1}^{n}\mathbb{E}\Bigg[\Bigg(\sum_{A\subsetneq[n]}k_{n,A}\sum_{j\notin A}\sqrt{\|\Delta_j f(X)\|^p + \|\tilde{\Delta}_i\Delta_j f(X)\|^p}$$

(5.27) $$\cdot \|\tilde{\Delta}_i\Delta_j f(X^A)\|\|\Delta_j f(X)\|\Bigg)^2\Bigg]$$

$$+ d^2\Big(\frac{9}{2} + \frac{9}{2p}\Big)\sum_{i=1}^{n}\mathbb{E}\Bigg[\Bigg(\sum_{A\subsetneq[n]}k_{n,A}\sum_{j\notin A}\sqrt{\|\Delta_j f(X)\|^p + \|\tilde{\Delta}_i\Delta_j f(X)\|^p}$$

$$\cdot \|\Delta_j f(X^A)\|\|\tilde{\Delta}_i\Delta_j f(X)\|\Bigg)^2\Bigg]$$

$$+ d^2\Big(\frac{9}{2} + \frac{9}{2p}\Big)\sum_{i=1}^{n}\mathbb{E}\Bigg[\Bigg(\sum_{A\subsetneq[n]}k_{n,A}\sum_{j\notin A}\sqrt{\|\Delta_j f(X)\|^p + \|\tilde{\Delta}_i\Delta_j f(X)\|^p}$$

$$\cdot \|\tilde{\Delta}_i\Delta_j f(X^A)\|\|\tilde{\Delta}_i\Delta_j f(X)\|\Bigg)^2\Bigg]$$

$$= d^2\gamma_{2+p}^{2+p}.$$

We also note that, using (5.24)

$$\mathbb{E}\Bigg[\Bigg(\sum_{A\subsetneq[n]}k_{n,A}\sum_{j\notin A}\mathbb{E}[|(\Delta_j f(X^A))_l||(\Delta_j f(X))_m||X]\Bigg)^2\Bigg]$$

(5.28) $$\leq \Bigg(\sum_{A\subsetneq[n]}k_{n,A}\sum_{j\notin A}\mathbb{E}[|(\Delta_j f(X^A))_l||(\Delta_j f(X))_m|]\Bigg)^2$$

$$+ \frac{1}{2} \sum_{i=1}^{n} \mathbb{E}\left[\left(\sum_{A \subsetneq [n]} k_{n,A} \sum_{j \notin A} \tilde{\Delta}_i\big(|(\Delta_j f(X^A))_l||(\Delta_j f(X))_m|\big)\right)^2\right]$$

$$< \infty,$$

by the assumption that $\mathbb{E}\|\Delta_j f(X)\|^4 < \infty$ for all $j \in [n]$. This shows that the variance in the definition of $V_{l,m}^{(1)}$ and $V_{l,m}^{(2)}$ exists.

5.2.6. *Step 5—concluding argument.* Using (5.13)–(5.22) and (5.27), we obtain that, for all $t \in (0, 1)$,

$$\big|\mathbb{E}\langle \nabla f_{t,h,\Sigma}(W), W\rangle - \langle \text{Hess}(f_{t,h,\Sigma}), T\rangle_{\text{H.S.}}\big|$$

(5.29)
$$\leq \frac{\sqrt{6}d^{3/2}}{2\sqrt{t}}\|\Sigma^{-1}\|_{\text{op}}^{3/2}\left(\sqrt{\sum_{k,l,m=1}^{d}(U_{k,l,m}^{(2)})^2} + \sqrt{d\sum_{l,m=1}^{d}(R_{l,m}^{(1)})^2} + \sqrt{d\sum_{l,m=1}^{d}(R_{l,m}^{(2)})^2}\right)$$

$$\leq \frac{\sqrt{6}d^{3/2}}{2\sqrt{t}}\|\Sigma^{-1}\|_{\text{op}}^{3/2}\Big(d^{3/2}\gamma_2^2 + d^2\|\Sigma^{-1/2}\|_{\text{op}}\gamma_2^2 + 2d_{\text{convex}}(W, N_\Sigma)d^{3/2}\gamma_1$$

$$+ d^{1/2}\sqrt{2 \cdot d_{\text{convex}}(W, N_\Sigma)d^2\gamma_3^3 + 2\|\Sigma^{-1/2}\|_{\text{op}}d^{5/2}\gamma_4^4}\Big).$$

Therefore, using (5.9), (5.11), (5.12) and (5.29),

$$d_{\text{convex}}(W, N_\Sigma)$$

$$\leq \frac{4}{3}\|\Sigma^{-1}\|_{\text{op}}d\sqrt{(\log t)^2 d_{\text{convex}}(W, N_\Sigma)} + 530d^{5/6}\sqrt{\mathbb{E}\|\mathbb{E}[T - \Sigma|W]\|_{\text{H.S.}}^2}$$

$$+ \frac{2\sqrt{6}d^{3/2}}{3\sqrt{t}}\|\Sigma^{-1}\|_{\text{op}}^{3/2}\Big(d^{3/2}\gamma_2^2 + d^2\|\Sigma^{-1/2}\|_{\text{op}}\gamma_2^2 + 2d_{\text{convex}}(W, N_\Sigma)d^{3/2}\gamma_1$$

$$+ d^{3/2}\sqrt{2 \cdot d_{\text{convex}}(W, N_\Sigma)\gamma_3^3 + 2d^{1/2}\|\Sigma^{-1/2}\|_{\text{op}}\gamma_4^4}\Big) + \frac{20}{\sqrt{2}}d\frac{\sqrt{t}}{1-t},$$

for all $t \in (0, 1)$. Assuming that $t \in (0, 1/2)$, we have $t^{1/4}|\log t| \leq 2$ and $1 - t \geq 1/2$ and so, in this case,

$$d_{\text{convex}}(W, N_\Sigma)$$

$$\leq \frac{4}{3}\|\Sigma^{-1}\|_{\text{op}}d\sqrt{\frac{4}{\sqrt{t}}d_{\text{convex}}(W, N_\Sigma)} + 530d^{5/6}\sqrt{\mathbb{E}\|\mathbb{E}[T - \Sigma|W]\|_{\text{H.S.}}^2}$$

$$+ \frac{2\sqrt{6}d^{3/2}}{3\sqrt{t}}\|\Sigma^{-1}\|_{\text{op}}^{3/2}\Big(d^{3/2}\gamma_2^2 + d^2\|\Sigma^{-1/2}\|_{\text{op}}\gamma_2^2 + 2d_{\text{convex}}(W, N_\Sigma)d^{3/2}\gamma_1$$

$$+ d^{3/2}\sqrt{2 \cdot d_{\text{convex}}(W, N_\Sigma)\gamma_3^3 + 2d^{1/2}\|\Sigma^{-1/2}\|_{\text{op}}\gamma_4^4}\Big) + \frac{40}{\sqrt{2}}d\sqrt{t}.$$

Suppose that $\gamma < \frac{1}{\sqrt{2}}$ (otherwise (2.4) clearly holds) and choose $\sqrt{t} = \max\{\frac{\sqrt{2}}{80d}d_{\text{convex}}(W, N_\Sigma), \gamma\}$. Then

$$d_{\text{convex}}(W, N_\Sigma) \leq \frac{4}{3}\|\Sigma^{-1}\|_{\text{op}}d\sqrt{\frac{320d}{\sqrt{2}} + 530d^{5/6}\gamma}$$

$$+ \frac{2\sqrt{6}d^{3/2}}{3}\|\Sigma^{-1}\|_{\text{op}}^{3/2}(d^{3/2}\gamma + d^2\|\Sigma^{-1/2}\|_{\text{op}}\gamma + 80\sqrt{2}d^{5/2}\gamma$$

$$+ d^{3/2}\sqrt{80\sqrt{2}d + 2d^{1/2}\|\Sigma^{-1/2}\|_{\text{op}}}\gamma) + \frac{40}{\sqrt{2}}d\gamma + \frac{1}{2}d_{\text{convex}}(W, N_\Sigma)$$

and (2.4) follows, as required.

5.3. *Proof of Lemma* 2.4. We will follow a strategy similar to that used in the proof of [24], Theorem 5.1. Recall that $T := \sum_{A \subsetneq [n]} k_{n,A} T_A/2$ for $k_{n,A} = \frac{1}{\binom{n}{|A|}(n-|A|)}$. Note that, using the fact that, for any $g : \mathcal{X}^{2n} \to \mathbb{R}$ and $U = g(X, X')$, $\text{Var}(\mathbb{E}(U|X)) \leq \mathbb{E}(\text{Var}(U|X'))$ (see [6], Lemma 4.4), we have

$$(5.30) \qquad \begin{aligned} \sqrt{\mathbb{E}\|\mathbb{E}[T|X] - \mathbb{E}T\|_{\text{H.S.}}^2} &\leq \frac{1}{2}\sum_{A \subsetneq [n]} k_{n,A}\sqrt{\mathbb{E}\|\mathbb{E}[T_A|X] - \mathbb{E}T_A\|_{\text{H.S.}}^2} \\ &\leq \frac{1}{2}\sum_{A \subsetneq [n]} k_{n,A}\sqrt{\mathbb{E}\|T_A - \mathbb{E}[T_A|X']\|_{\text{H.S.}}^2}. \end{aligned}$$

Introduce the substitution operator

$$\tilde{S}_i(X) = (X_0, \dots, \tilde{X}_i, \dots, X_n).$$

Fix $A \subsetneq [n]$ and note that, by the Efron–Stein inequality [48],

$$(5.31) \qquad \frac{1}{2}\sum_{A \subsetneq [n]} k_{n,A}\sqrt{\mathbb{E}\|T_A - \mathbb{E}[T_A|X']\|_{\text{H.S.}}^2} \leq \frac{1}{\sqrt{8}}\sum_{A \subsetneq [n]} k_{n,A}\sqrt{\sum_{i=1}^{n}\mathbb{E}\|\tilde{\Delta}_i T_A\|_{\text{H.S.}}^2}.$$

Recall also that $T_A = \sum_{j \notin A}[\Delta_j f(X)][\Delta_j f(X^A)]^T \in \mathbb{R}^{d \times d}$ and so

$$(5.32) \qquad \begin{aligned} &\sum_{i=1}^{n}\mathbb{E}\|\tilde{\Delta}_i T_A\|_{\text{H.S.}}^2 \\ &= \sum_{i=1}^{n}\sum_{j,k \notin A}\mathbb{E}\langle\tilde{\Delta}_i(\Delta_j f(X)\Delta_j f(X^A)^T), \tilde{\Delta}_i(\Delta_k f(X)\Delta_k f(X^A)^T)\rangle_{\text{H.S.}}. \end{aligned}$$

Moreover, note that for $j \notin A$ and a fixed $i \in [n]$,

$$(5.33) \quad \tilde{\Delta}_i(\Delta_j f(X)\Delta_j f(X^A)^T) = \tilde{\Delta}_i\Delta_j f(X)\Delta_j f(X^A)^T + \Delta_j f(\tilde{S}_i(X))\tilde{\Delta}_i\Delta_j f(X^A)^T.$$

We shall now analyse each summand in (5.32) separately. In order to do this, we introduce the vector $\bar{X}$ by $\bar{X}_i = \tilde{X}_i$ and $\bar{X}_l = X'_l$, if $l \neq i$ and for any $x \in \mathcal{X}^n$ and any mapping $\varphi : \mathcal{X}^n \to \mathbb{R}^d$, we define

$$\bar{\Delta}_l\varphi(x) = \varphi(x) - \varphi(x_1, \dots, x_{l-1}, \bar{X}_l, x_{l+1}, \dots, x_{l+1}, \dots, x_n).$$

Note that, if i, j, k are pairwise distinct, then, by (5.33), the corresponding summands in (5.32) are upper-bounded by

$$\begin{aligned} &4\sup_{(Y,Y',Z,Z')}\mathbb{E}|\langle\bar{\Delta}_i\bar{\Delta}_j f(Y)\bar{\Delta}_j f(Y')^T, \bar{\Delta}_i\bar{\Delta}_k f(Z)\bar{\Delta}_k f(Z')^T\rangle_{\text{H.S.}}| \\ &= 4\sup_{(Y,Y',Z,Z')}\mathbb{E}|\langle\Delta_i\Delta_j f(Y)\Delta_j f(Y')^T, \Delta_i\Delta_k f(Z)\Delta_k f(Z')^T\rangle_{\text{H.S.}}| \\ &\leq 4\sup_{(Y,Y',Z,Z')}\mathbb{E}[\mathbb{1}_{\{\Delta_{i,j} f(Y)\neq 0, \Delta_{i,k} f(Z)\neq 0\}}(\|\Delta_j f(Y)\Delta_j f(Y')^T\|_{\text{H.S.}} \\ &\quad + \|\Delta_j f(Y^i)\Delta_j f(Y')^T\|_{\text{H.S.}}) \end{aligned}$$

$$\cdot \big(\big\|\Delta_k f(Y)\Delta_k f(Y')^T\big\|_{\text{H.S.}} + \big\|\Delta_k f(Y^i)\Delta_k f(Y')^T\big\|_{\text{H.S.}}\big)\big]$$

(5.34)

$$= 4 \sup_{(Y,Y',Z,Z')} \mathbb{E}\big[\mathbb{1}_{\{\Delta_{i,j}f(Y)\neq 0,\Delta_{i,k}f(Z)\neq 0\}}\big(\big\|\Delta_j f(Y)\big\|\big\|\Delta_j f(Y')\big\|$$

$$+ \big\|\Delta_j f(Y^i)\big\|\big\|\Delta_j f(Y')\big\|\big)$$

$$\cdot \big(\big\|\Delta_k f(Y)\big\|\big\|\Delta_k f(Y')\big\| + \big\|\Delta_k f(Y^i)\big\|\big\|\Delta_k f(Y')\big\|\big)\big]$$

$$\leq 16 \sup_{(Y,Y',Z,Z')} \mathbb{E}\big[\mathbb{1}_{\{\Delta_{i,j}f(Y)\neq 0,\Delta_{i,k}f(Y')\neq 0\}}\big\|\Delta_j f(Z)\big\|^2\big\|\Delta_k f(Z')\big\|^2\big]$$

$$= 16B_n',$$

where Y, Y', Z, Z' are recombinations of $\{X, X', \tilde{X}\}$. Now, using similar computations and the Cauchy–Schwarz inequality, we obtain that, in the case $i \neq j = k$, the corresponding summands in (5.32) are upper-bounded by

$$4 \sup_{(Y,Y',Z,Z')} \mathbb{E}\big|\langle \bar{\Delta}_i \bar{\Delta}_j f(Y)\bar{\Delta}_j f(Y')^T, \bar{\Delta}_i \bar{\Delta}_k f(Z)\bar{\Delta}_k f(Z')^T\rangle_{\text{H.S.}}\big|$$

$$\leq 4 \sup_{(Y,Y',Z,Z')} \mathbb{E}\big[\big\|\bar{\Delta}_i \bar{\Delta}_j f(Y)\big\|\big\|\bar{\Delta}_j f(Y')\big\|\big\|\bar{\Delta}_i \bar{\Delta}_k f(Z)\big\|\big\|\bar{\Delta}_k f(Z')\big\|\big]$$

(5.35)

$$\leq 4 \sup_{(Y,Y')} \mathbb{E}\big[\big\|\Delta_j \Delta_i f(Y)\big\|^2\big\|\Delta_j f(Y')\big\|^2\big]$$

$$\leq 4 \sup_{(Y,Y')} \mathbb{E}\big[\mathbb{1}_{\{\Delta_{i,j}f(Y)\neq 0\}}\big(\big\|\Delta_i f(Y)\big\| + \big\|\Delta_i f(Y^j)\big\|\big)^2\big\|\Delta_j f(Y')\big\|^2\big]$$

$$\leq 16 \sup_{(Y,Y',Z)} \mathbb{E}\big[\mathbb{1}_{\{\Delta_{i,j}f(Y)\neq 0\}}\big\|\Delta_i f(Z)\big\|^2\big\|\Delta_j f(Y')\big\|^2\big] = 16B_n(f).$$

We note that, in the case $i = j$, using $\tilde{\Delta}_i \Delta_i(\cdot) = \tilde{\Delta}_i(\cdot)$, the Hilbert–Schmidt norm of the right-hand side of (5.33) can be bounded by

$$\big\|\tilde{\Delta}_i f(X)\Delta_i f(X^A)^T\big\|_{\text{H.S.}} + \big\|\Delta_i f(\tilde{S}_i(X))\tilde{\Delta}_i f(X^A)^T\big\|_{\text{H.S.}}$$

(5.36)

$$= \big\|\tilde{\Delta}_i f(X)\big\|\big\|\Delta_i f(X^A)\big\| + \big\|\Delta_i f(\tilde{S}_i(X))\big\|\big\|\tilde{\Delta}_i f(X^A)\big\|$$

$$\leq \frac{1}{2}\big[\big\|\tilde{\Delta}_i f(X)\big\|^2 + \big\|\Delta_i f(X^A)\big\|^2 + \big\|\Delta_i f(\tilde{S}_i(X))\big\|^2 + \big\|\tilde{\Delta}_i f(X^A)\big\|^2\big].$$

If $i \notin A$ and $i = j = k$ then it follows from (5.36), that the corresponding summands in (5.32) are smaller than

(5.37)

$$\frac{1}{4}\mathbb{E}\big[\big(\big\|\Delta_i f(X)\big\|^2 + \big\|\tilde{\Delta}_i f(X^A)\big\|^2 + \big\|\Delta_i f(\tilde{S}_i(X))\big\|^2 + \big\|\tilde{\Delta}_i f(X^A)\big\|^2\big)^2\big]$$

$$\leq 4\mathbb{E}\big\|\Delta_i f(X)\big\|^4.$$

Finally, the corresponding summands in (5.32) in the case $i = j \neq k$ are upper-bounded by

$$4 \sup_{(Y,Y',Z)} \mathbb{E}\big|\langle \bar{\Delta}_i f(Y)^T \bar{\Delta}_i f(Y), \bar{\Delta}_i \bar{\Delta}_k f(Y')^T \bar{\Delta}_k f(Z)\rangle_{\text{H.S.}}\big|$$

(5.38)

$$\leq 4 \sup_{(Y,Y',Z)} \mathbb{E}\big[\mathbb{1}_{\{\Delta_{i,k}f(Y)\neq 0\}}\big\|\Delta_i f(Y')\big\|^2\big(\big\|\Delta_k f(Y)\big\| + \big\|\Delta_k f(Y^i)\big\|\big)\big\|\Delta_k f(Z)\big\|\big]$$

$$\leq 8B_n(f).$$

Therefore, by (5.30)–(5.32), (5.34), (5.35), (5.37) and (5.38),

$$
\sqrt{\mathbb{E}\big\|\mathbb{E}[T\,|\,X] - \mathbb{E}T\big\|_{\mathrm{H.S.}}^2} \leq \sqrt{2n} \sum_{A\subsetneq[n]} k_{n,A}\bigg(\sum_{j,k\notin A} \big[\mathbb{1}_{[j=k=1]}\mathbb{E}\big\|\Delta_1 f(X)\big\|^4
$$

$$
+ (\mathbb{1}_{[j=k\neq1]} + \mathbb{1}_{[k\neq j=1]})B_n(f) + \mathbb{1}_{[1\neq k\neq j\neq1]}B_n'(f)\big]\bigg)^{1/2}
$$

$$
\leq \sqrt{2n}\bigg(\sqrt{\mathbb{E}\big\|\Delta_1 f(X)\big\|^4} \sum_{A\subsetneq[n]:1\notin A} k_{n,A}
$$

$$
+ \sqrt{2B_n(f)} \sum_{A\subsetneq[n]} k_{n,A}\sqrt{n-|A|} + \sqrt{B_n'(f)} \sum_{A\subsetneq[n]} k_{n,A}\big(n-|A|\big)\bigg)
$$

$$
\leq 4\sqrt{n}\Big(\sqrt{nB_n(f)} + \sqrt{n^2 B_n'(f)} + \sqrt{\mathbb{E}\big\|\Delta_1 f(X)\big\|^4}\Big),
$$

where we used the fact that $\sum_{A\subsetneq[n]} k_{n,A}\sqrt{n-|A|} = \sum_{k=1}^n \frac{1}{\sqrt{k}} \leq \int_0^n \frac{dx}{\sqrt{x}} = 2\sqrt{n}$.

6. Proofs of the results related to covering processes. In the arguments below we shall use the following notation. For any $k \in [n]$ and a set of k pairwise distinct indices $T = \{i_1, \ldots, i_k\} \subset [n]$ and for any $Z = (Z_1, \ldots, Z_n)$ such that $Z_l \subset \mathbb{R}^d$ for all $l \in [n]$,

$$
F_n^T(Z) = F_n^{(i_1, i_2, \ldots, i_k)}(Z) := \bigcup_{\substack{1\leq l\leq n \\ l\notin\{i_1,\ldots,i_k\}}} Z_l.
$$

We will also often write $F_n^T := F_n^T(X)$ and $F_n^{(i_1, i_2, \ldots, i_k)} := F_n^{(i_1, i_2, \ldots, i_k)}(X)$. For such Z and for any $1 \leq i \neq j \leq n$, let $Z^{\{j\}} = (Z_1, \ldots, Z_{j-1}, X_j', Z_{j+1}, \ldots, Z_n)$, $Z_{\{i\}} = (Z_1, \ldots, Z_{i-1}, \tilde{X}_i, Z_{i+1}, \ldots, Z_n)$ and

$$
Z^{\{i\},\{j\}} = \big(Z_1^{\{i\},\{j\}}, \ldots, Z_n^{\{i\},\{j\}}\big),
$$

where, for $l \in [n]$, $Z_l^{\{i\},\{j\}} = Z_l^{\{j\}}$, if $l \neq i$ and $Z_l^{\{i\},\{j\}} = \tilde{X}_i$, if $l = i$. Furthermore, for $i \neq j$ we define

$$
\Delta_j f(Z) = f(Z) - f(Z^{\{j\}}), \qquad \tilde{\Delta}_i \Delta_j f(Z) = f(Z) - f(Z^{\{j\}}) - f(Z_{\{i\}}) + f(Z^{\{i\},\{j\}}).
$$

By $\mathcal{K}^d$ we will denote the set of compact convex subsets of $\mathbb{R}^d$. We will also often use the well-known bound $\min(\kappa_0, \ldots, \kappa(d)) \geq \frac{1}{d!}$, which readily gives that $V_i(L) \leq d!\overline{V}(L)$, for any $i = 0, \ldots, d$ and $L \in \mathcal{K}^d$. Finally, we will repeatedly use the fact that, for any $i = 0, \ldots, d$, V_i is a translation invariant and additive functional on the ring of finite unions of convex bodies in $\mathbb{R}^d$ and it is nonnegative and monotone on $\mathcal{K}^d$.

6.1. *Initial lemmas.* We start by quoting a result from [26].

LEMMA 6.1 (Proposition 22.5 in [26]). *For all $L, M \in \mathcal{K}^d$,*

$$
\int_{\mathbb{R}^d} \overline{V}\big((x+L)\cap M\big)\,dx \leq \overline{V}(L)\overline{V}(M).
$$

Now, we prove a useful estimate.

LEMMA 6.2. *For all $L \in \mathcal{K}^d$, $k = 0, \ldots, n-1$, and $i = 0, \ldots, d$,*

(A) $\quad \mathbb{E}\left| \overline{V}\left(L \cap \bigcup_{l=k+1}^{n} X_l \right) \right| \leq e^{\overline{V}(K)} \overline{V}(L);$

(B) $\quad \mathbb{E}\left| \overline{V}\left(L \cap \bigcup_{l=k+1}^{n} X_l \right) \right|^m \leq (2^{2^d + 2d} d^{d/2})^m e^{2^m (2R+1)^d} \overline{V}(L)^m \quad \text{for all } m \geq 1.$

PROOF. To prove (A), we apply the inclusion-exclusion formula and the triangle inequality:

$$\mathbb{E}\left| \overline{V}\left(L \cap \bigcup_{l=k+1}^{n} X_l \right) \right| \leq \sum_{l=1}^{n} \binom{n}{l} \mathbb{E}[\overline{V}(L \cap X_1 \cap \cdots \cap X_l)]$$

$$\leq \sum_{l=1}^{n} \frac{\binom{n}{l}}{n^l} \int_{E_n} \cdots \int_{E_n} \overline{V}(L \cap (K + x_1) \cap \cdots \cap (K + x_l)) \, dx_1 \cdots dx_l$$

$$\overset{\text{Lemma 6.1}}{\leq} \overline{V}(L) \sum_{l=1}^{n} \frac{1}{l!} \overline{V}(K)^l \leq e^{\overline{V}(K)} \overline{V}(L).$$

To prove (B), we use a strategy similar to that of the proof of [26], Proposition 22.4. Let $Q_0 = [-1/2, 1/2]^d$ and, for any $z \in \mathbb{R}^d$, $Q_z = Q_0 + z$. Let $I(L) = \{z \in \mathbb{Z}^d : Q_z \cap L \neq \varnothing\}$ and, for $C \subset \mathbb{R}^d$, let $N(C) := \sum_{l=k+1}^{n} \mathbb{1}_{[X_l \cap C \neq \varnothing]}$ be the number of grains in $\{X_{k+1}, \ldots, X_n\}$ hitting C. For each nonempty $I \subset I(L)$, we fix some $z(I) \in I$. As in the proof of [26], Proposition 22.4, the inclusion-exclusion formula yields

(6.1) $\quad \mathbb{E}\left| V_i\left(L \cap \bigcup_{l=k+1}^{n} X_l \right) \right|^m \leq \mathbb{E}\left| \sum_{I \subset I(L): I \neq \varnothing} \mathbb{1}\left[\bigcap_{z \in I} Q_z \neq \varnothing \right] 2^{N(Q_{z(I)})} \overline{V}(Q_0) \right|^m.$

Using [26], (22.26), and the display directly below it, we obtain that

(6.2) $\quad \text{card}\left\{ I \subset I(L) : I \neq \varnothing, \bigcap_{z \in I} Q_z \neq \varnothing \right\} \leq 2^{2^d} \sum_{l=0}^{d} \kappa_{d-l} d^{(d-l)/2} V_i(L) \leq 2^{2^d} d^{d/2} \overline{V}(L).$

Now, for any $z \in \mathbb{R}^d$, let $Q_{R,z} = [z - 1/2 - R, z + 1/2 + R]^d$. For any $l = k+1, \ldots, n$, let U_l be the center of X_l. Note that, for n large enough so that $Q_{R,z} \subset E_n$,

$$\mathbb{E}[2^{mN(Q_z)}] \leq \mathbb{E}[2^{m \sum_{l=k+1}^{n} \mathbb{1}_{[U_l \in Q_{R,z}]}}] = (\mathbb{E}[2^{m \mathbb{1}_{[U_1 \in Q_{R,z}]}}])^{n-k}$$

(6.3) $$\leq (1 + 2^m \mathbb{E}[\mathbb{1}_{[U_1 \in Q_{R,z}]}])^n \leq \left(1 + 2^m \frac{V_d(Q_{R,z})}{n} \right)^n$$

$$\leq \exp(2^m (2R+1)^d).$$

The result now follows from (6.1)–(6.3) and the fact that $\overline{V}(Q_0) = \sum_{l=0}^{d} \kappa_{d-l} \binom{d}{l} \leq \sum_{l=0}^{d} (\sqrt{2\pi})^{d-l} \binom{d}{l} \leq 4^d$ (see [28], Example 1.3, and [43], page 224–227). $\quad \square$

6.2. *Proof of Theorem 3.1.*

6.2.1. *Introduction.*
Let η be the stationary Poisson process with the intensity measure $\Lambda(\cdot) = \int \mathbb{1}_{[K+x \in \cdot]} dx$ and let $Z \equiv Z(\eta)$ be given by $Z := \bigcup_{L \in \eta} L$. For $i = 0, \ldots, d$, let

$V_i^*(L) := \mathbb{E}V_i(Z \cap L) - V_i(L)$, for $L \in \mathcal{K}^d$. It follows from [22], Theorem 6.4, (5.2), Theorem 3.1, that

$$\Sigma_{i,j} = \sum_{k=2}^{\infty} \frac{1}{k!} \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} V_i^*(K \cap (K + x_2) \cap \cdots \cap (K + x_k))$$
$$\cdot V_j^*(K \cap (K + x_2) \cap \cdots \cap (K + x_k)) \, dx_2 \cdots dx_k$$

and the infinite sum in the definition of Σ converges by [22], Theorem 3.1. Therefore, Σ is well defined.

Now, for every vector $a = (a_0, \ldots, a_d)^T \in \mathbb{R}^{d+1}$,

$$a^T \Sigma a = \sum_{k=2}^{\infty} \frac{1}{k!} \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} \left(\sum_{l=0}^{d} a_l V_l^*(K \cap (K + x_2) \cap \cdots \cap (K + x_k)) \right)^2 dx_2 \cdots dx_k.$$

In order to prove positive-definiteness of Σ, it suffices to show that the summand obtained for $k = d + 1$ is strictly positive. This is shown in the proof of [22], Theorem 4.1, so we readily obtain that Σ is positive-definite.

Next, we prove lemmas that will help us bound the distance between Σ_n and Σ.

6.2.2. *Auxiliary lemmas.* Now, we prove the first part of the bound on the rate of convergence to the limiting covariance.

LEMMA 6.3. *For $i = 0, \ldots, d$, $k = 1, 2, \ldots$, let*
$$\psi_k^i(x_1, \ldots, x_k) := \mathbb{E}[V_i(F_n^{(1,2,\ldots,k)} \cap x_1 \cap \cdots \cap x_k)] - V_i(x_1 \cap \cdots \cap x_k).$$
If $n > \overline{V}(K)e$, then

$$\left| \mathrm{Cov}((f_i(X), f_j(X)) - \frac{1}{n} \sum_{k=1}^{n} \binom{n}{k} \mathbb{E}[\psi_k^i(X_1, \ldots, X_k)\psi_k^j(X_1, \ldots, X_k)] \right.$$
$$\left. + \mathbb{E}[\psi_1^i(X_1)]\mathbb{E}[\psi_1^j(X_1)] \right|$$
$$\le (d!)^2 (1 + e^{\overline{V}(K)})^2 (\overline{V}(K)^2 e^{\overline{V}(K) + 3\overline{V}(K)^2} + e^{\overline{V}(K)^2}) n^{-1}.$$

PROOF. Let $[k] = \{1, \ldots, k\}$ and for all $i = 0, \ldots, d$ and $k = 1, \ldots, n$, let

$$(6.4) \qquad \phi_k^i(x_1, \ldots, x_k) = \frac{1}{\sqrt{n}} \sum_{m=0}^{k} (-1)^m \sum_{\substack{T \subset [k] \\ |T| = m}} \varphi_T^i(x_1, \ldots, x_k),$$

where, for $T \subset [k]$, such that $|T| = m$ and $T = \{t_1, \ldots, t_m\}$,

$$\varphi_T^i(x_1, \ldots, x_k) := \mathbb{E}[V_i(x_{t_1} \cap x_{t_2} \cap \cdots \cap x_{t_m} \cap X_1 \cap \cdots \cap X_{k-m})]$$
$$- \mathbb{E}[V_i(F_n^{(1,2,\ldots,k)} \cap x_{t_1} \cap x_{t_2} \cap \cdots \cap x_{t_m} \cap X_1 \cap \cdots \cap X_{k-m})].$$

Note that $\psi_k^i(x_1, \ldots, x_k) = -\varphi_{[k]}^i(x_1, \ldots, x_k)$. Using [24], Theorem 2.2, we have that

$$\mathrm{Cov}(f_i(X), f_j(X))$$

$$(6.5) \qquad = \sum_{k,l=1}^{n} \sum_{1 \le r_1 < \cdots < r_k \le n} \sum_{1 \le s_1 < \cdots < s_l \le n} (-1)^{k+l} \mathbb{E}[\phi_k^i(X_{r_1}, \ldots, X_{r_k})\phi_l^j(X_{s_1}, \ldots, X_{s_l})]$$

$$= \sum_{k=1}^{n} \binom{n}{k} \mathbb{E}[\phi_k^i(X_1, \ldots, X_k)\phi_k^j(X_1, \ldots, X_k)].$$

More precisely, identity (6.5) follows from the fact that, for $L \in \mathcal{K}^d$ and for all $i = 0, \ldots, d$ and $k, j_1, \ldots, j_k = 1, \ldots, n$,

$$V_i\big((F_n^{(j_1,\ldots,j_k)} \cup X_{j_k}) \cap L\big) - V_i\big((F_n^{(j_1,\ldots,j_k)} \cup X'_{j_k}) \cap L)\big)$$
$$= V_i(X_{j_k} \cap L) - V_i\big(F_n^{(j_1,\ldots,j_k)} \cap X_{j_k} \cap L\big) - V_i(X'_{j_k} \cap L) + V_i\big(F_n^{(j_1,\ldots,j_k)} \cap X'_{j_k} \cap L\big)$$

and, as a consequence, using the notation of [24],

$$\Delta_1 \Delta_2 \cdots \Delta_k f_i(X)$$
$$= \frac{(-1)^{k+1}}{\sqrt{n}}\big[V_i(X_1 \cap \cdots \cap X_k) - V_i\big(F_n^{(1,2,\ldots,k)} \cap X_1 \cap \cdots \cap X_k\big)\big]$$
$$+ \sum_{m=1}^{k} \frac{(-1)^{m+k+1}}{\sqrt{n}} \sum_{1 \le t_1 < \cdots < t_m \le k} \big[V_i\big(X'_{t_1} \cap X'_{t_2} \cap \cdots \cap X'_{t_m} \cap X_1 \cap \cdots \cap X_{k-m}\big)$$
$$- V_i\big(F_n^{(1,2,\ldots,k)} \cap X'_{t_1} \cap X'_{t_2} \cap \cdots \cap X'_{t_m} \cap X_1 \cap \cdots \cap X_{k-m}\big)\big].$$

Now, for $m_1, m_2 = 0, \ldots, k$, $m_1 \ge m_2$, $i, j = 0, \ldots, d$, $T = \{t_1, \ldots, t_{m_1}\} \subset [k]$, $S = \{s_1, \ldots, s_{m_2}\} \subset [k]$, such that $|T| = m_1$, $|S| = m_2$,

$$\frac{1}{n}\big|\mathbb{E}\varphi_T^i(X_1, \ldots, X_k)\varphi_S^j(X_1, \ldots, X_k)\big|$$
$$= \frac{1}{n}\big|\mathbb{E}\big\{\big[V_i\big(X_{t_1} \cap \cdots \cap X_{t_{m_1}} \cap X'_1 \cap \cdots \cap X'_{k-m_1}\big)$$
$$- V_i\big(F_n^{(1,2,\ldots,k)}(X') \cap X_1 \cap \cdots \cap X_{t_{m_1}} \cap X'_1 \cap \cdots \cap X'_{k-m_1}\big)\big]$$
$$\cdot \big[V_j\big(X_{s_1} \cap \cdots \cap X_{s_{m_2}} \cap \tilde{X}_1 \cap \cdots \cap \tilde{X}_{k-m_2}\big)$$
$$- V_j\big(F_n^{(1,2,\ldots,k)}(\tilde{X}) \cap X_{s_1} \cap \cdots \cap X_{s_{m_2}} \cap \tilde{X}_1 \cap \cdots \cap \tilde{X}_{k-m_2}\big)\big]\big\}\big|$$
$$\overset{\text{Lemma 6.2}}{\le} \frac{(d!)^2(1 + e^{\overline{V}(K)})^2}{n}\mathbb{E}\big[\overline{V}\big(X_{t_1} \cap \cdots \cap X_{t_{m_1}} \cap X'_1 \cap \cdots \cap X'_{k-m_1}\big)$$
$$\cdot \overline{V}\big(X_{s_1} \cap \cdots \cap X_{s_{m_2}} \cap \tilde{X}_1 \cap \cdots \cap \tilde{X}_{k-m_2}\big)\big]$$
$$\overset{\text{Lemma 6.1}}{\le} (d!)^2(1 + e^{\overline{V}(K)})^2$$
$$\cdot \Big(\frac{1}{n^{2k-1}}\overline{V}(K)^{2k}\mathbb{1}_{[T \cap S = \varnothing]} + \frac{1}{n^{2k-|T \cap S|}}\big(\overline{V}(K)^2\big)^{k-|T \cap S|/2+1}\mathbb{1}_{[T \cap S \ne \varnothing]}\Big).$$

Now, we consider the case where $T \cap S \ne \varnothing$ and ($T \ne [k]$ or $S \ne [k]$). In the expression below, $l = |T \cap S|$. Note that, using the above calculation and Lemma 6.1,

$$\frac{1}{n}\sum_{k=1}^{n}\binom{n}{k}\sum_{m_1,m_2=0}^{k} \sum_{\substack{T, S \subset [k] \\ |S \cap T| \notin \{0,k\} \\ |T|=m_1, |S|=m_2}} \big|\mathbb{E}\varphi_T^i(X_1, \ldots, X_k)\varphi_S^j(X_1, \ldots, X_k)\big|$$
$$\le (d!)^2(1 + e^{\overline{V}(K)})^2 \sum_{k=1}^{n}\binom{n}{k}\sum_{l=1}^{k-1}\binom{k}{l}\sum_{r_1=0}^{k-l}\binom{k-l}{r_1}$$
$$\cdot \sum_{r_2=0}^{k-l-r_1}\binom{k-l-r_1}{r_2}\frac{(\overline{V}(K)^2)^{k-l/2+1}}{n^{2k-l}}$$

$$(6.6) \qquad = (d!)^2(1 + e^{\overline{V}(K)})^2 \sum_{k=1}^{n} \binom{n}{k} \sum_{l=1}^{k-1} \binom{k}{l} \sum_{r_1=0}^{k-l} \binom{k-l}{r_1} 2^{k-l-r_1} \frac{(\overline{V}(K)^2)^{k-l/2+1}}{n^{2k-l}}$$

$$= (d!)^2(1 + e^{\overline{V}(K)})^2 \sum_{k=1}^{n} \binom{n}{k} \sum_{l=1}^{k-1} \binom{k}{l} 3^{k-l} \frac{(\overline{V}(K)^2)^{k-l/2+1}}{n^{2k-l}}$$

$$\leq (d!)^2(1 + e^{\overline{V}(K)})^2 \overline{V}(K)^2 \sum_{k=1}^{n} \frac{n^k}{k!}(3\overline{V}(K)^2)^k \sum_{l=1}^{k-1} \binom{k}{l} \frac{1}{n^{2k-l}} (3\overline{V}(K))^{-l}$$

$$\leq \frac{(d!)^2(1 + e^{\overline{V}(K)})^2 \overline{V}(K)^2}{n} \sum_{k=1}^{n} \frac{1}{k!}(3\overline{V}(K)^2)^k ((3\overline{V}(K))^{-1} + 1)^k$$

$$\leq (d!)^2 \overline{V}(K)^2 (1 + e^{\overline{V}(K)})^2 e^{\overline{V}(K)+3\overline{V}(K)^2} n^{-1}.$$

Also, note that, in the case $T \cap S = \varnothing$, for $k \geq 2$,

$$\frac{1}{n} \sum_{k=2}^{n} \binom{n}{k} \Big| \sum_{\substack{T,S \subset [k] \\ T \cap S = \varnothing}} \mathbb{E}[\varphi_T^i(X_1, \ldots, X_k)\varphi_S^j(X_1, \ldots, X_k)] \Big|$$

$$= \frac{1}{n} \sum_{k=2}^{n} \binom{n}{k} \Big| \sum_{\substack{T,S \subset [k] \\ T \cap S = \varnothing}} (-1)^{|T|+|S|} \mathbb{E}[V_i(F_n^{(1,2,\ldots,k)} \cap X_1 \cap \cdots \cap X_k)$$

$$- V_i(X_1 \cap \cdots \cap X_k)]$$

$$\cdot \mathbb{E}[V_j(F_n^{(1,2,\ldots,k)} \cap X_1 \cap \cdots \cap X_k) - V_j(X_1 \cap \cdots \cap X_k)] \Big|$$

$$= \frac{1}{n} \sum_{k=2}^{n} \binom{n}{k} \Big| \sum_{t=0}^{k} (-1)^t \binom{k}{t} \sum_{s=0}^{k-t} (-1)^s \binom{k-t}{s} \mathbb{E}[V_i(F_n^{(1,2,\ldots,k)} \cap X_1 \cap \cdots \cap X_k)$$

$$- V_i(X_1 \cap \cdots \cap X_k)] \cdot \mathbb{E}[V_j(F_n^{(1,2,\ldots,k)} \cap X_1 \cap \cdots \cap X_k)$$

$$- V_j(X_1 \cap \cdots \cap X_k)] \Big|$$

$$(6.7) \qquad = \frac{1}{n} \sum_{k=2}^{n} \binom{n}{k} \Big| \sum_{t=0}^{k} (-1)^t \binom{k}{t} \mathbb{1}[k=t] \mathbb{E}[V_i(F_n^{(1,2,\ldots,k)} \cap X_1 \cap \cdots \cap X_k)$$

$$- V_i(X_1 \cap \cdots \cap X_k)]$$

$$\cdot \mathbb{E}[V_j(F_n^{(1,2,\ldots,k)} \cap X_1 \cap \cdots \cap X_k) - V_j(X_1 \cap \cdots \cap X_k)] \Big|$$

$$= \frac{1}{n} \sum_{k=2}^{n} \binom{n}{k} \Big| (-1)^k \mathbb{E}[V_i(F_n^{(1,2,\ldots,k)} \cap X_1 \cap \cdots \cap X_k \cap E_n)$$

$$- V_i(X_1 \cap \cdots \cap X_k)]$$

$$\cdot \mathbb{E}[V_j(F_n^{(1,2,\ldots,k)} \cap X_1 \cap \cdots \cap X_k) - V_j(X_1 \cap \cdots \cap X_k)] \Big|$$

$$\leq \sum_{k=2}^{n} \frac{(d!)^2(1 + e^{\overline{V}(K)})^2}{n^{2k-1}} \binom{n}{k} \overline{V}(K)^{2k}$$

$$\leq (d!)^2 \left(1 + e^{\overline{V}(K)}\right)^2 \sum_{k=2}^{n} \frac{\overline{V}(K)^{2k}}{n^{k-1}k!}$$

$$\leq (d!)^2 \left(1 + e^{\overline{V}(K)}\right)^2 e^{\overline{V}(K)^2} n^{-1}$$

where we use Lemmas 6.1 and 6.2 in the last three inequalities. The result now follows from (6.4)–(6.7). □

Now, we show the second part of the bound on the rate of convergence to the limiting covariance.

LEMMA 6.4. *Using the notation of Lemma* 6.3, *we have that, for all* $i, j = 0, \ldots, d$ *and* Σ *defined in* (3.1),

$$\left| \frac{1}{n} \sum_{k=1}^{n} \binom{n}{k} \mathbb{E}[\psi_k^i(X_1, \ldots, X_k) \psi_k^j(X_1, \ldots, X_k)] - \mathbb{E}[\psi_1^i(X_1)] \mathbb{E}[\psi_1^j(X_1)] - \Sigma_{i,j} \right|$$

$$\leq (19 + 8(4R+1)^d)(d!)^2 \max(1, \overline{V}(K)^3)(e^{\overline{V}(K)} + 1)^2 e^{\max(1, 2\overline{V}(K))} n^{-1/d}.$$

PROOF. Let $E_{(n^{1/d}+4R)^d} \subseteq \mathbb{R}^d$ be the cube with volume $(n^{1/d} + 4R)^d$, centered at the origin. Note that, for any $l \in \mathbb{N}$, $i = 0, \ldots, d$,

$$\frac{1}{n} \left| \int_{E_n} \cdots \int_{E_n} V_i((K + x_1) \cap \cdots \cap (K + x_l)) dx_1 \cdots dx_l \right.$$

$$\left. - \int_{E_n} \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} V_i((K + x_1) \cap \cdots \cap (K + x_l)) dx_1 \cdots dx_l \right|$$

$$\leq \frac{1}{n} \sum_{m=1}^{l-1} \underbrace{\int_{E_n} \cdots \int_{E_n}}_{m \text{ times}} \underbrace{\int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d}}_{(l-m-1) \text{ times}} \int_{\mathbb{R}^d \backslash E_n} V_i((K + x_1) \cap \cdots \cap (K + x_l)) dx_1 \cdots dx_l$$

(6.8)
$$\leq \frac{l-1}{n} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} \int_{E_{(n^{1/d}+4R)^d} \backslash E_n} V_i((K + x_1) \cap \cdots \cap (K + x_l)) dx_1 \cdots dx_l$$

$$\stackrel{(*)}{=} \frac{l-1}{n} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} V_i((K + x_1) \cap \cdots \cap (K + x_{l-1}) \cap K) dx_1 \cdots dx_{l-1}$$

$$\cdot \int_{E_{(n^{1/d}+4R)^d} \backslash E_n} dx_l$$

$$\stackrel{(**)}{\leq} \frac{(l-1)d!}{n} \left((n^{1/d} + 4R)^d - n\right) \overline{V}(K)^l$$

$$\leq \frac{(l-1)(d!)(4R+1)^d}{n^{1/d}} \overline{V}(K)^l,$$

where we have used translation invariance of V_i and of the Lebesgue measure on $\mathbb{R}^d$ in $(*)$ and Lemma 6.1 in $(**)$.

Now, we use Fubini's theorem and Hadwiger's general integral geometric theorem [44], Theorem 5.1.2, in a manner analogous to the one in which it was done in the first display on [44], page 387. We note that K stays invariant under rotations about the origin. We further use the formula of the second display on [44], page 388 (which is derived from

the Crofton formula [44], (5.6)). In total, letting $c_j^m = \frac{m!\kappa_m}{j!\kappa_j}$, $m, j \in \{0, \ldots, d\}$ and $(V_i)_m = c_i^{d-m} c_d^{m+i} V_{m+i} \mathbb{1}_{[m+i \leq d]}$, $i, m \in \{0, \ldots, d\}$, we obtain that

$$\int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} V_i\big((K + x_1) \cap \cdots \cap (K + x_l) \cap L\big)\, dx_1 \cdots dx_l$$

$$= \sum_{\substack{0 \leq r_0, \ldots, r_l \leq d \\ r_0 + \cdots + r_l = ld}} c_{d-r_0}^d \big[(V_i)_{r_0}(L)\big] \prod_{m=1}^{l} c_d^{r_m} V_{r_m}(K)$$

(6.9)

$$= \sum_{r_0=0}^{d} \sum_{\substack{0 \leq r_1, \ldots, r_l \leq d \\ r_1 + \cdots + r_l = ld - r_0}} c_{d-r_0}^d c_i^{d-r_0} c_d^{r_0+i} V_{r_0+i}(L) \mathbb{1}_{[r_0+i \leq d]} \prod_{m=1}^{l} c_d^{r_m} V_{r_m}(K)$$

$$\overset{s := r_0 + i}{=} \sum_{s=i}^{d} \tilde{P}_{s,l,i}(d) V_s(L).$$

for

$$\tilde{P}_{s,l,i}(d) = \sum_{\substack{0 \leq r_1, \ldots, r_l \leq d \\ r_1 + \cdots + r_l = ld - s + i}} \frac{s!\kappa_s}{i!\kappa_i} \prod_{m=1}^{l} \frac{r_m!\kappa_{r_m}}{d!\kappa_d} V_{r_m}(K).$$

Furthermore, using the inclusion-exclusion principle, for any $L \in \mathcal{K}^d$,

$$\mathbb{E}\big[V_i\big(F_n^{(1,\ldots,k)} \cap L\big)\big]$$

(6.10)

$$= \sum_{l=1}^{n-k} \frac{\binom{n-k}{l}}{n^l} (-1)^{l-1} \int_{E_n} \cdots \int_{E_n} V_i\big((K + x_1) \cap \cdots \cap (K + x_l) \cap L\big)\, dx_1 \cdots dx_l.$$

It follows that

$$\Bigg| \frac{1}{n} \int_{E_n} \cdots \int_{E_n} \Big[\mathbb{E} V_i\big(F_n^{(1,\ldots,k)} \cap (K + x_1) \cap \cdots \cap (K + x_k)\big)$$

$$- V_i\big((K + x_1) \cap \cdots \cap (K + x_k)\big)\Big]$$

$$\cdot \Big[\mathbb{E} V_j\big(F_n^{(1,\ldots,k)} \cap (K + x_1) \cap \cdots \cap (K + x_k)\big)$$

$$- V_j\big((K + x_1) \cap \cdots \cap (K + x_k)\big)\Big] dx_1 \cdots dx_k$$

$$- \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} \Bigg[\sum_{l=1}^{n-k} \frac{(-1)^{l-1} \binom{n-k}{l}}{n^l}$$

$$\cdot \sum_{s=i}^{d} \tilde{P}_{s,l,i}(d) V_s\big(K \cap (K + x_2) \cap \cdots \cap (K + x_k)\big)$$

$$- V_i\big(K \cap (K + x_2) \cdots \cap (K + x_k)\big)\Bigg]$$

$$\cdot \Bigg[\sum_{l=1}^{n-k} \frac{(-1)^{l-1} \binom{n-k}{l}}{n^l} \sum_{s=j}^{d} \tilde{P}_{s,l,j}(d) V_s\big(K \cap (K + x_2) \cap \cdots \cap (K + x_k)\big)$$

$$- V_j\big(K \cap (K + x_2) \cap \cdots \cap (K + x_k)\big)\Bigg] dx_2 \cdots dx_k \Bigg|$$

$$= \frac{1}{n}\left|\int_{E_n}\cdots\int_{E_n}\left[\mathbb{E}V_i\big(F_n^{(1,\ldots,k)}\cap(K+x_1)\cap\cdots\cap(K+x_k)\big)\right.\right.$$
$$\left.- V_i\big((K+x_1)\cap\cdots\cap(K+x_k)\big)\right]$$
(6.11)
$$\cdot\left[\mathbb{E}V_j\big(F_n^{(1,\ldots,k)}\cap(K+x_1)\cap\cdots\cap(K+x_k)\big)\right.$$
$$\left.- V_j\big((K+x_1)\cap\cdots\cap(K+x_k)\big)\right]dx_1\cdots dx_k$$
$$-\int_{E_n}\int_{\mathbb{R}^d}\cdots\int_{\mathbb{R}^d}\left[\sum_{l=1}^{n-k}\frac{(-1)^{l-1}\binom{n-k}{l}}{n^l}\right.$$
$$\cdot\sum_{s=i}^{d}\tilde{P}_{s,l,i}(d)V_s\big((K+x_1)\cap\cdots\cap(K+x_k)\big)$$
$$\left.- V_i\big((K+x_1)\cap\cdots\cap(K+x_k)\big)\right]$$
$$\cdot\left[\sum_{l=1}^{n-k}\frac{(-1)^{l-1}\binom{n-k}{l}}{n^l}\sum_{s=j}^{d}\tilde{P}_{s,l,j}(d)V_s\big((K+x_1)\cap\cdots\cap(K+x_k)\big)\right.$$
$$\left.\left.- V_j\big((K+x_1)\cap\cdots\cap(K+x_k)\big)\right]dx_1\cdots dx_k\right|$$
$$\leq 2d!\left(\sum_{l=1}^{n-k}\frac{\binom{n-k}{l}}{n^l}\overline{V}(K)^{l+1}+\overline{V}(K)\right)$$
$$\cdot\left(\sum_{l=1}^{n-k}\frac{\binom{n-k}{l}}{n^l}\frac{(l+k-1)(d!)(4R+1)^d}{n^{1/d}}\overline{V}(K)^{k+l}\right.$$
$$\left.+\frac{(k-1)(d!)(4R+1)^d}{n^{1/d}}\overline{V}(K)^k\right)$$
$$\leq 2(d!)^2(4R+1)^d\overline{V}(K)^{k+1}\big(e^{\overline{V}(K)}+1\big)\big((\overline{V}(K)+k-1)e^{\overline{V}(K)}+k-1\big)n^{-1/d},$$

where the first identity holds by the translation invariance of V_i and the Lebesgue measure on $\mathbb{R}^d$ and the second to last inequality follows from (6.10), (6.9), (6.8) and Lemma 6.1.

Now, we use the trick applied in the last display on [44], page 387, in the derivation of [44], Theorem 9.1.3. In the calculation below, t will denote the number of indices $r_1,\ldots,r_l$ which are smaller than d and $r=l-t$. We obtain a result analogous to that of the third display on [44], page 388, namely

$$\sum_{l=1}^{\infty}\frac{(-1)^{l-1}}{l!}\sum_{s=i}^{d}\tilde{P}_{s,l,i}(d)V_s(L)-V_i(L)$$
$$=\sum_{l=1}^{\infty}\frac{(-1)^{l-1}}{l!}\sum_{s=i}^{d}V_s(L)\sum_{\substack{0\leq r_1,\ldots,r_l\leq d\\ r_1+\cdots+r_l=ld-s+i}}\frac{s!\kappa_s}{i!\kappa_i}\prod_{m=1}^{l}\frac{r_m!\kappa_{r_m}}{d!\kappa_d}V_{r_m}(K)-V_i(L)$$
$$=\sum_{s=i}^{d}\sum_{l=1}^{\infty}\frac{(-1)^{l-1}}{l!}V_i(L)\left[\mathbb{1}_{[s=i]}V_d(L)^l+\mathbb{1}_{[s\neq i]}V_s(L)\sum_{t=1}^{s-i}\binom{l}{t}V_d(K)^{l-t}\right.$$

$$(6.12) \qquad \cdot \sum_{\substack{0 \le r_1, \ldots, r_t \le d-1 \\ r_1 + \cdots + r_t = td - s + i}} \frac{s! \kappa_s}{i! \kappa_i} \prod_{m=1}^{t} \frac{r_m! \kappa_{r_m}}{d! \kappa_d} V_{r_m}(K) \Bigg] - V_i(L)$$

$$= -e^{-V_d(K)} V_i(L) - \mathbb{1}_{[i \ne d]} \sum_{s=i+1}^{d} V_s(L) \sum_{r=0}^{\infty} V_d(K)^r \sum_{t=1}^{s-i} \frac{(-1)^r}{(r+t)!} \binom{r+t}{t} (-1)^t$$

$$\cdot \sum_{\substack{0 \le r_1, \ldots, r_t \le d-1 \\ r_1 + \cdots + r_t = td - s + i}} \frac{s! \kappa_s}{i! \kappa_i} \prod_{m=1}^{t} \frac{r_m! \kappa_{r_m}}{d! \kappa_d} V_{r_m}(K)$$

$$= -\sum_{s=i}^{d} V_s(L) P_{i,s}(d),$$

for

$$P_{i,s}(d) = e^{-V_d(K)} \Bigg[\mathbb{1}_{[s=i]} + \mathbb{1}_{[i \ne d]} \frac{s! \kappa_s}{i! \kappa_i} \sum_{t=1}^{s-i} \frac{(-1)^t}{t!} \sum_{\substack{i \le r_1, \ldots, r_t \le d-1 \\ r_1 + \cdots + r_t = td + i - s}} \prod_{m=1}^{t} \frac{r_m! \kappa_{r_m}}{d! \kappa_d} V_{r_m}(K) \Bigg].$$

It follows that

$$\Bigg| \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} \Bigg[\sum_{l=1}^{n-k} \frac{(-1)^{l-1} \binom{n-k}{l}}{n^l} \sum_{s=i}^{d} \tilde{P}_{s,l,i}(d) V_s\big(K \cap (K + x_2) \cap \cdots \cap (K + x_k)\big)$$

$$- V_i\big(K \cap (K + x_2) \cap \cdots \cap (K + x_k)\big) \Bigg]$$

$$\cdot \Bigg[\sum_{l=1}^{n-k} \frac{(-1)^{l-1} \binom{n-k}{l}}{n^l} \sum_{s=j}^{d} \tilde{P}_{s,l,j}(d) V_s\big(K \cap (K + x_2) \cap \cdots \cap (K + x_k)\big)$$

$$- V_j\big(K \cap (K + x_2) \cap \cdots \cap (K + x_k)\big) \Bigg] dx_2 \cdots dx_k$$

$$- \sum_{s=i}^{d} \sum_{r=j}^{d} \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} P_{i,s}(d) P_{j,r}(d) V_s\big(K \cap (K + x_2) \cdots \cap (K + x_k)\big)$$

$$\cdot V_r\big(K \cap (K + x_2) \cdots \cap (K + x_k)\big) dx_2 \cdots dx_k \Bigg|$$

$$\overset{(\dagger)}{\le} (d!)^2 \Bigg[\sum_{l=1}^{\infty} \frac{\overline{V}(K)^{l+1}}{l!} + \sum_{l=1}^{n-k} \frac{\binom{n-k}{l}}{n^l} \overline{V}(K)^{l+1} + 2\overline{V}(K) \Bigg]$$

$$(6.13) \qquad \cdot \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} \Bigg[\Bigg(\sum_{l=0}^{\infty} \frac{\overline{V}(K)^l}{l!}$$

$$- \sum_{l=0}^{n-k} \overline{V}(K)^l \frac{\binom{n-k}{l}}{n^l} \Bigg) \overline{V}\big(K \cap (K + x_2) \cap \cdots \cap (K + x_k)\big) \Bigg] dx_2 \cdots dx_k$$

$$\overset{(\ddagger)}{\le} 2(d!)^2 \overline{V}(K)^{k+1} \big(e^{\overline{V}(K)} + 1\big) \Bigg(e^{\overline{V}(K)} - \Big(1 + \frac{\overline{V}(K)}{n}\Big)^{n-k} \Bigg)$$

$$\leq 2(d!)^2 \overline{V}(K)^{k+1}(e^{\overline{V}(K)} + 1)\left[\left(e^{\overline{V}(K)} - \left(1 + \frac{\overline{V}(K)}{n}\right)^n\right)\right.$$

$$+ \left.\left(\left(1 + \frac{\overline{V}(K)}{n}\right)^n - \left(1 + \frac{\overline{V}(K)}{n}\right)^{n-k}\right)\right]$$

$$\leq 2(d!)^2 \overline{V}(K)^{k+1}(e^{\overline{V}(K)} + 1)\left[\sum_{l=n+1}^{\infty} \frac{\overline{V}(K)^l}{l!} + \sum_{r=1}^{n}\left(\frac{\overline{V}(K)^r}{r!} - \frac{\overline{V}(K)^r \binom{n}{r}}{n^r}\right)\right.$$

$$+ \left.\sum_{m=1}^{k}\left(\left(1 + \frac{\overline{V}(K)}{n}\right)^{n-m+1} - \left(1 + \frac{\overline{V}(K)}{n}\right)^{n-m}\right)\right]$$

$$\overset{(\P)}{\leq} 2(d!)^2 \overline{V}(K)^{k+2}(e^{\overline{V}(K)} + 1)\left[\frac{e}{n} + \frac{2e^{2\overline{V}(K)}}{n} + k\frac{e^{\overline{V}(K)}}{n}\right],$$

where inequality (†) follows from (6.9), (6.12) and Lemma 6.1; inequality (‡) follows from Lemma 6.1; and inequality (¶) follows from the well-known bound $\sum_{k=B}^{\infty} \frac{A^k}{k!} \leq (\frac{Ae}{B})^B$, for $B > A > 0$, our assumption $n > \overline{V}(K)e$ and the fact that, for $r, n \in \mathbb{N} \setminus \{0\}$, $r \leq n$,

(6.14)
$$\left|\frac{1}{r!} - \frac{\binom{n}{r}}{n^r}\right| = \left|\frac{n^r - n(n-1)\cdots(n-r+1)}{n^r r!}\right| \leq \frac{n^r - (n-r+1)^r}{n^r r!}$$

$$\leq \frac{1}{r!}\sum_{l=0}^{r-1}\binom{r}{l}\frac{r^{r-l}}{n^{r-l}} \leq \frac{1}{(r-1)!n}\sum_{l=0}^{r-1}\binom{r}{l} \leq \frac{2^r}{n(r-1)!}.$$

Now, using (6.14), (6.9), (6.12), (6.10) and Lemmas 6.1 and 6.2,

(6.15)
$$\sum_{k=1}^{n}\left|\frac{\binom{n}{k}}{n^k} - \frac{1}{k!}\right|\left|\int_{\mathbb{R}^d}\cdots\int_{\mathbb{R}^d}\sum_{s=i}^{d}\sum_{r=j}^{d} P_{i,s}(d)P_{j,r}(d)V_s\big(K \cap (K + x_2) \cap \cdots \cap (K + x_k)\big)\right.$$

$$\cdot V_r\big(K \cap (K + x_2) \cap \cdots \cap (K + x_k)\big)\,dx_2\cdots dx_k$$

$$\leq \sum_{k=1}^{n}\frac{2^k}{(k-1)!}(d!)^2 \overline{V}(K)^{k+1}(e^{\overline{V}(K)} + 1)^2\frac{1}{n}$$

$$\leq 2\overline{V}(K)^2(e^{\overline{V}(K)} + 1)^2(d!)^2 e^{2\overline{V}(K)}n^{-1}.$$

Similarly,

(6.16)
$$\sum_{k=n+1}^{\infty}\frac{1}{k!}\int_{\mathbb{R}^d}\cdots\int_{\mathbb{R}^d}\sum_{s=i}^{d}\sum_{r=j}^{d} P_{i,s}(d)P_{j,r}(d)V_s\big(K \cap (K + x_2) \cap \cdots \cap (K + x_k)\big)$$

$$\cdot V_r\big(K \cap (K + x_2) \cap \cdots \cap (K + x_k)\big)\,dx_2\cdots dx_k$$

$$\leq \sum_{k=n+1}^{\infty}\frac{1}{k!}(d!)^2 \overline{V}(K)^{k+1}(e^{\overline{V}(K)} + 1)^2 \leq \frac{\overline{V}(K)^2 e(d!)^2}{n}(e^{\overline{V}(K)} + 1)^2.$$

Using (6.11), (6.13), (6.15), (6.16), we obtain

$$
\begin{aligned}
&\left|\frac{1}{n}\sum_{k=1}^{n}\binom{n}{k}\mathbb{E}\big[\psi_k^i(X_1,\ldots,X_k)\psi_k^j(X_1,\ldots,X_k)\big]\right.\\
&\qquad -\sum_{k=1}^{\infty}\frac{1}{k!}\sum_{s=i}^{d}\sum_{r=j}^{d}\int_{\mathbb{R}^d}\cdots\int_{\mathbb{R}^d}P_{i,s}(d)P_{j,r}(d)\\
&\qquad\cdot V_s\big(K\cap(K+x_2)\cap\cdots\cap(K+x_k)\big)\\
&\qquad\cdot V_r\big(K\cap(K+x_2)\cap\cdots\cap(K+x_k)\big)\,dx_2\cdots dx_k\Bigg|\\
&\leq (d!)^2(e^{\overline{V}(K)}+1)^2\\
&\quad\cdot\big(e^{2\overline{V}(K)}(6\overline{V}(K)^2+2\overline{V}(K)^3)+3\overline{V}(K)^2e+6(4R+1)^d\overline{V}(K)^2e^{\overline{V}(K)}\big)n^{-1/d}.
\end{aligned}
\tag{6.17}
$$

Now, using (6.8), Lemma 6.1, and the translation invariance of the Lebesgue measure, we obtain

$$
\begin{aligned}
&\Bigg|\Bigg[\sum_{l=1}^{n-1}\frac{(-1)^l}{n^{l+1}}\binom{n-1}{l}\\
&\qquad\cdot\int_{E_n}\cdots\int_{E_n}V_i\big((K+x_1)\cap\cdots(K+x_{l+1})\big)\,dx_1\cdots dx_{l+1}-V_i(K)\Bigg]\\
&\qquad\cdot\Bigg[\sum_{l=1}^{n-1}\frac{(-1)^l}{n^{l+1}}\binom{n-1}{l}\\
&\qquad\cdot\int_{E_n}\cdots\int_{E_n}V_i\big((K+x_1)\cap\cdots(K+x_{l+1})\big)\,dx_1\cdots dx_{l+1}-V_i(K)\Bigg]\\
&\qquad-\Bigg[\sum_{l=1}^{n-1}\frac{(-1)^l}{n^l}\binom{n-1}{l}\\
&\qquad\cdot\int_{\mathbb{R}^d}\cdots\int_{\mathbb{R}^d}V_i\big(K\cap(K+x_2)\cap\cdots(K+x_{l+1})\big)\,dx_2\cdots dx_{l+1}-V_i(K)\Bigg]\\
&\qquad\cdot\Bigg[\sum_{l=1}^{n-1}\frac{(-1)^l}{n^l}\binom{n-1}{l}\\
&\qquad\cdot\int_{\mathbb{R}^d}\cdots\int_{\mathbb{R}^d}V_i\big(K\cap(K+x_2)\cap\cdots(K+x_{l+1})\big)\,dx_2\cdots dx_{l+1}-V_i(K)\Bigg]\Bigg|\\
&\leq 2(d!)\Bigg[\sum_{l=1}^{n-1}\frac{\binom{n-1}{l}}{n^l}\overline{V}(K)^{l+1}+\overline{V}(K)\Bigg]\Bigg[\sum_{l=1}^{n-1}\frac{\binom{n-1}{l}}{n^l}\frac{l(d!)(4R+1)^d}{n^{1/d}}\overline{V}(K)^l\Bigg]\\
&\leq 2(d!)^2\overline{V}(K)^2(e^{\overline{V}(K)}+1)(4R+1)^de^{\overline{V}(K)}n^{-1/d}.
\end{aligned}
\tag{6.18}
$$

Similarly, using (6.9), (6.12) and Lemma 6.1, in a manner similar to the calculation (6.13),

(6.19)
$$\left| \left[\sum_{l=1}^{n-1} \frac{(-1)^l}{n^{l+1}} \binom{n-1}{l} \right. \right.$$
$$\cdot \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} V_i(K \cap (K + x_2) \cap \cdots \cap (K + x_{l+1})) \, dx_2 \cdots dx_{l+1} - V_i(K) \right]$$
$$\cdot \left[\sum_{l=1}^{n-1} \frac{(-1)^l}{n^l} \binom{n-1}{l} \right.$$
$$\cdot \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} V_i(K \cap (K + x_2) \cap \cdots \cap (K + x_{l+1})) \, dx_2 \cdots dx_{l+1} - V_i(K) \right]$$
$$- \sum_{s=i}^{d} \sum_{r=j}^{d} P_{i,s}(d) P_{j,r}(d) V_s(K) V_r(K) \right|$$
$$\leq 2(d!)^2 \overline{V}(K)^3 (e^{\overline{V}(K)} + 1) \frac{e + 2e^{2\overline{V}(K)} + e^{\overline{V}(K)}}{n}.$$

From (6.18) and (6.19) we obtain that

(6.20)
$$\left| \mathbb{E}[\psi_1^i(X_1)] \mathbb{E}[\psi_1^j(X_1)] - \sum_{s=i}^{d} \sum_{r=j}^{d} P_{i,s}(d) P_{j,r}(d) V_s(K) V_r(K) \right|$$
$$\leq 2(d!)^2 \overline{V}(K)^2 (e^{\overline{V}(K)} + 1)$$
$$\cdot [(4R + 1)^d e^{\overline{V}(K)} + \overline{V}(K)(e + 2e^{2\overline{V}(K)} + e^{\overline{V}(K)})] n^{-1/d}$$

and the final bound follows from (6.17) and (6.20) and the fact that $\overline{V}(K) = \sum_{j=0}^{d} \binom{d}{j} \kappa_d R^j \leq 3^d (R+1)^d$. $\quad\square$

6.2.3. *Concluding argument.* The proof of Theorem 3.1 is finished by applying the discussion of Section 6.2.1, a direct application of Lemmas 6.3 and 6.4 and the fact that $\overline{V}(K) = \sum_{j=0}^{d} \binom{d}{j} \kappa_d R^j \leq 3^d (R+1)^d$ (see [28], Example 1.2, and [43], page 224–227).

6.3. *Proof of Theorem 3.2.* We will say that $Y = (Y_1, \ldots, Y_n)$ is a recombination of X, X', $\tilde{X}$ if $Y_i \in \{X_i, X_i', \tilde{X}_i\}$, for $i \in [n]$. In what follows, $Y, \bar{Y}, \hat{Y}, \tilde{Y}, Y', \bar{Y}', \hat{Y}', \tilde{Y}', Z, \bar{Z}, \hat{Z}, Z', \bar{Z}', \hat{Z}'$ will denote recombinations of X, X', $\tilde{X}$ and all the suprema in this section will be taken over all such recombinations.

6.3.1. *Auxiliary lemma.* First, we provide a useful estimate, which we prove in section A of the *supplement* [23].

LEMMA 6.5. *For all* $m, m_1, m_2, m_3 \geq 1$, *we have*

$$\mathbb{E}[\|\Delta_1 f(X)\|^m] \leq \frac{C_m^{(1)}(d)}{n^{m/2}};$$

$$\sup_{(Y, Z)} \mathbb{E}[\overline{V}(Y_1 \cap Y_2) \|\Delta_1 f(Z)\|^m] \leq \frac{C_m^{(2)}(d)}{n^{m/2+1}};$$

$$\sup_{(Y,Z)} \mathbb{E}\big[\|\Delta_2 f(Y)\|^{m_1}\|\tilde{\Delta}_1\Delta_2 f(Z)\|^{m_2}\big] \le \frac{C_{m_1,m_2}^{(3)}(d)}{n^{(m_1+m_2)/2+1}};$$

$$\sup_{(Y,Y',Z,Z')} \mathbb{E}\big[\|\Delta_2 f(Y)\|^{m_1}\|\Delta_3 f(Y')\|^{m_1}\|\tilde{\Delta}_1\Delta_2 f(Z)\|^{m_2}\|\tilde{\Delta}_1\Delta_3 f(Z')\|^{m_2}\big] \le \frac{C_{m_1,m_2}^{(4)}(d)}{n^{m_1+m_2+2}};$$

$$\sup_{(Y,Y',Z,Z')} \mathbb{E}\big[\overline{V}(Y_1 \cap Y_2)\overline{V}(Z_1 \cap Z_3)\|\Delta_2 f(Y')\|^m\|\Delta_3 f(Z')\|^m\big] \le \frac{C_m^{(5)}(d)}{n^{m+2}},$$

$$\sup_{(Y,Y',Z)} \mathbb{E}\big[\|\tilde{\Delta}_1 f(Y)\|^{m_1}\|\Delta_2 f(Y')\|^{m_2}\|\tilde{\Delta}_1\Delta_2 f(Z)\|^{m_3}\big] \le \frac{C_{m_1,m_2,m_3}^{(6)}(d)}{n^{(m_1+m_2+m_3)/2+1}};$$

$$\sup_{(Y,Y',Z)} \mathbb{E}\big[\|\Delta_1 f(Y)\|^{m_1}\|\Delta_2 f(Y')\|^{m_2}\|\tilde{\Delta}_1\Delta_2 f(Z)\|^{m_3}\big] \le \frac{C_{m_1,m_2,m_3}^{(6)}(d)}{n^{(m_1+m_2+m_3)/2+1}};$$

$$\sup_{(Y,Y',Z)} \mathbb{E}\big[\overline{V}(Y_1 \cap Y_2)\|\Delta_1 f(Y')\|^m\|\Delta_2 f(Z)\|^m\big] \le \frac{C_{2m}^{(2)}(d)}{n^{m+1}},$$

for

$$C_m^{(1)}(d) := (2d+4)^m (d!)^m \big(2^{2^d} d^{d/2} 12^d (R+1)^d\big)^m e^{2^m (2R+1)^d};$$

$$C_m^{(2)}(d) := (4d+8)^m (d!)^m \big(2^{2^d+2d} d^{d/2}\big)^m \big(3(R+1)\big)^{d(m+2)} e^{2^m (2R+1)^d};$$

$$C_{m_1,m_2}^{(3)}(d) := (4d+8)^{m_1+m_2} (d!)^{m_1+m_2} \big(2^{2^d+2d} d^{d/2}\big)^{m_1+m_2}$$
$$\cdot \big(3(R+1)\big)^{m_1+m_2+1} e^{(2^{2m_1-1}+2^{2m_2-1})(2R+1)^d};$$

$$C_{m_1,m_2}^{(4)}(d) := (8d+16)^{2m_1+2m_2} (d!)^{2m_1+2m_2} \big(2^{2^d} d^{d/2} 12^d (R+1)^d\big)^{2m_1+2m_2}$$
$$\cdot \big(3(R+1)\big)^{2d} e^{(2^{4m_1-1}+2^{4m_2-1})(2R+1)^d};$$

$$C_m^{(5)}(d) := (8d+16)^{2m} (d!)^{2m} \big(2^{2^d+2d} d^{d/2}\big)^{2m} \big(3(R+1)\big)^{d(2m+4)} e^{2^{2m}(2R+1)^d};$$

$$C_{m_1,m_2,m_3}^{(6)}(d) := (4d+8)^{m_1+m_2+m_3} (d!)^{m_1+m_2+m_3} \big(2^{2^d} d^{d/2} 12^d (R+1)^d\big)^{m_1+m_2+m_3}$$
$$\cdot \big(3(R+1)\big)^d e^{\frac{1}{3}(2^{3m_1}+2^{3m_2}+2^{3m_3})(2R+1)^d}.$$

6.3.2. *Concluding argument.* Using Lemma 6.5, and adopting the notation of Theorem 2.2 we have that

$$\gamma_1 \le \frac{C_3^{(1)}(d)}{n^{1/2}} \quad \text{and} \quad \gamma_2 \le \frac{\sqrt{C_4^{(1)}(d)}}{n^{1/2}}$$

and so

(6.21) $$\max(\gamma_1, \gamma_2) \le 8(d+2)^3 (d!)^3 \big(2^{2^d+2d} d^{d/2}\big)^3 \big(3(R+1)\big)^{3d} e^{8(2R+1)^d} n^{-1/2}.$$

Now, we look at γ_3 and γ_4 of Theorem 2.2. For $A \subsetneq [n]$, let $k_{n,A} := \frac{1}{\binom{n}{|A|}(n-|A|)}$. We have that, for $p = 1, 2$,

$$\gamma_{p+2}^{p+2} = \sum_{i=1}^n \sum_{A_1,A_2 \subsetneq [n]} k_{n,A_1} k_{n,A_2} \sum_{j \notin A_1} \sum_{k \notin A_2} \Big[\frac{3}{2}\beta_{n,A_1,A_2}^{(1,p)}(i,j,k)$$
(6.22)
$$+ \Big(\frac{9}{2} + \frac{9}{2p}\Big)\big(\beta_{n,A_1,A_2}^{(2,p)}(i,j,k) + \beta_{n,A_1,A_2}^{(3,p)}(i,j,k) + \beta_{n,A_1,A_2}^{(4,p)}(i,j,k)\big)\Big].$$

where, for $A_1, A_2 \subsetneq [n]$, $i, j, k \in [n]$ such that $j \notin A_1$, $k \notin A_2$, and $p = 1, 2$,

$$
\begin{aligned}
\beta_{n,A_1,A_2}^{(1,p)} &:= \mathbb{E}\big\{ \mathbb{1}_{[\tilde{\Delta}_i \Delta_j f(X) \neq 0]} \mathbb{1}_{[\tilde{\Delta}_i \Delta_k f(X) \neq 0]} \sqrt{\|\Delta_j f(X)\|^p + \|\tilde{\Delta}_i \Delta_j f(X)\|^p} \\
&\quad \cdot \sqrt{\|\Delta_k f(X)\|^p + \|\tilde{\Delta}_i \Delta_k f(X)\|^p} \\
&\quad \cdot \|\Delta_j f(X^{A_1})\| \|\Delta_k f(X^{A_2})\| \|\Delta_j f(X)\| \|\Delta_k f(X)\| \big\},
\end{aligned}
$$

$$
\begin{aligned}
\beta_{n,A_1,A_2}^{(2,p)} &:= \mathbb{E}\big\{ \sqrt{\|\Delta_j f(X)\|^p + \|\tilde{\Delta}_i \Delta_j f(X)\|^p} \cdot \sqrt{\|\Delta_k f(X)\|^p + \|\tilde{\Delta}_i \Delta_k f(X)\|^p} \\
&\quad \cdot \|\tilde{\Delta}_i \Delta_j f(X^{A_1})\| \|\Delta_j f(X)\| \|\tilde{\Delta}_i \Delta_k f(X^{A_2})\| \|\Delta_k f(X)\| \big\},
\end{aligned}
$$

$$
\begin{aligned}
\beta_{n,A_1,A_2}^{(3,p)} &:= \mathbb{E}\big\{ \sqrt{\|\Delta_j f(X)\|^p + \|\tilde{\Delta}_i \Delta_j f(X)\|^p} \cdot \sqrt{\|\Delta_k f(X)\|^p + \|\tilde{\Delta}_i \Delta_k f(X)\|^p} \\
&\quad \cdot \|\tilde{\Delta}_i \Delta_j f(X)\| \|\Delta_j f(X^{A_1})\| \|\tilde{\Delta}_i \Delta_k f(X)\| \|\Delta_k f(X^{A_2})\| \big\},
\end{aligned}
$$

$$
\begin{aligned}
\beta_{n,A_1,A_2}^{(4,p)} &:= \mathbb{E}\big\{ \sqrt{\|\Delta_j f(X)\|^p + \|\tilde{\Delta}_i \Delta_j f(X)\|^p} \cdot \sqrt{\|\Delta_k f(X)\|^p + \|\tilde{\Delta}_i \Delta_k f(X)\|^p} \\
&\quad \cdot \|\tilde{\Delta}_i \Delta_j f(X^{A_1})\| \|\tilde{\Delta}_i \Delta_j f(X)\| \|\tilde{\Delta}_i \Delta_k f(X^{A_2})\| \|\tilde{\Delta}_i \Delta_k f(X)\| \big\}.
\end{aligned}
$$

Let us adopt the notation of Lemma 6.5. We have the following four lemmas, which are proved in section B of the *supplement* [23].

LEMMA 6.6.   *For $j = k = i \notin A_1 \cup A_2$,*

$$
\beta_{n,A_1,A_2}^{(l,p)}(i,j,k) \leq \frac{2C_{p+4}^{(1)}(d)}{n^{p/2+2}} \quad \text{for all } l = 1, \ldots, 4.
$$

LEMMA 6.7.   *For ($i = j \neq k$ or $i = k \neq j$), $j \notin A_1$, $k \notin A_2$,*

$$
\beta_{n,A_1,A_2}^{(1,p)}(i,j,k) \leq 8(2^{p/2} + 1)(d!) \frac{C_{p+4}^{(2)}(d)}{n^{p/2+3}};
$$

$$
\beta_{n,A_1,A_2}^{(4,p)}(i,j,k) \leq 2(2^{p/2} + 1) \frac{C_{p/2,p/2,4}^{(6)}(d)}{n^{p/2+3}};
$$

$$
\beta_{n,A_1,A_2}^{(2,p)}(i,j,k) + \beta_{n,A_1,A_2}^{(3,p)}(i,j,k) \leq 4(2^{p/2} + 1) \frac{\sqrt{C_{2,p/2+1,1}^{(6)}(d) C_{p+2,p/2+1,1}^{(6)}(d)}}{n^{p/2+3}}.
$$

LEMMA 6.8.   *For $i \neq k = j$, $j \notin A_1 \cup A_2$,*

$$
\beta_{n,A_1,A_2}^{(1,p)}(i,j,k) \leq \frac{4(2^p + 1)d! C_{p+4}^{(2)}(d)}{n^{p/2+3}};
$$

$$
\beta_{n,A_1,A_2}^{(2,p)}(i,j,k) + \beta_{n,A_1,A_2}^{(3,p)}(i,j,k) \leq \frac{2(C_{2,p+2}^{(3)}(d) + C_{p+2,2}^{(3)}(d))}{n^{p/2+3}};
$$

$$
\beta_{n,A_1,A_2}^{(4,p)}(i,j,k) \leq \frac{(2^p + 1)C_{p,4}^{(3)}(d)}{n^{p/2+3}}.
$$

LEMMA 6.9.   *If i, j, k are pairwise distinct and $k \notin A_2$, $j \notin A_1$,*

$$
\beta_{n,A_1,A_2}^{(1,p)}(i,j,k) \leq \frac{16(2^{p/2} + 1)^2 d! C_{p/2+2}^{(5)}(d)}{n^{p/2+4}};
$$

$$\beta_{n,A_1,A_2}^{(2,p)}(i,j,k) + \beta_{n,A_1,A_2}^{(3,p)}(i,j,k) \leq \frac{2(2^{p/2}+1)^2 C_{p/2+1,1}^{(4)}(d)}{n^{p/2+4}};$$

$$\beta_{n,A_1,A_2}^{(4,p)}(i,j,k) \leq \frac{(2^{p/2}+1)^2 C_{p/2,2}^{(4)}(d)}{n^{p/2+4}}.$$

It now follows from (6.22) and Lemmas 6.6–6.9 that, for $p = 1, 2$,

$$
\begin{aligned}
\gamma_{p+2} \leq \Big\{ &\frac{3}{2}\big[2C_{p+4}^{(1)}(d) + (8(2^{p/2}+1) + 4(2^p+1))(d!)C_{p+2}^{(2)}(d) \\
&+ 16(2^{p/2}+1)^2(d!)C_{p/2+2}^{(5)}(d)\big] \\
&+ \Big(\frac{9}{2} + \frac{9}{2p}\Big)\big[6C_{p+4}^{(1)}(d) + 4(2^{p/2}+1)\sqrt{C_{2,p/2+1,1}^{(6)}(d)C_{p+2,p/2+1,1}^{(6)}(d)} \\
&+ 2(2^{p/2}+1)C_{p/2,p/2,4}^{(6)}(d) \\
&+ 2(C_{2,p+2}^{(3)}(d) + C_{p+2,p}^{(3)}(d) + (2^p+1)C_{p,4}^{(3)}(d)) \\
&+ 2(2^{p/2}+1)^2 C_{p/2+1,1}^{(4)}(d) + (2^{p/2}+1)^2 C_{p/2,2}^{(4)}(d)\big]\Big\}^{1/(p+2)} n^{-1/2} \\
\leq\ &215(d+2)^{5/3}(d!)^{5/3}(2^{2^d+2d}d^{d/2})^{5/3}(3(R+1))^{5d/3}e^{456(2R+1)^d}n^{-1/2}.
\end{aligned}
$$
(6.23)

Now, note that each V_l is an additive functional on the ring of finite unions of convex bodies. Using the notation introduced in Section 2.2, it follows that

$$
\begin{aligned}
\mathbb{1}_{[\Delta_{1,2}f(Y)\neq 0]} &\leq \mathbb{1}_{[Y_1\cap Y_2\neq\varnothing]} + \mathbb{1}_{[Y_1\cap X_2'\neq\varnothing]} + \mathbb{1}_{[X_1'\cap Y_2\neq\varnothing]} + \mathbb{1}_{[X_1'\cap X_2'\neq\varnothing]} \\
&\leq d!\big[\overline{V}(Y_1\cap Y_2) + \overline{V}(Y_1\cap X_2') + \overline{V}(X_1'\cap Y_2) + \overline{V}(X_1'\cap X_2')\big],
\end{aligned}
$$

for any recombination Y of $(X, X', \tilde{X})$. Together with Lemmas 2.4 and 6.5 and using the notation theoreof, we therefore obtain

$$B_n(f) \leq 4d! \sup_{(Y,Z,Z')} \mathbb{E}\big[\overline{V}(Y_1\cap Y_2)\|\Delta_1 f(Z)\|^2\|\Delta_2 f(Z')\|^2\big] \leq \frac{4d!C_4^{(2)}(d)}{n^3};$$

$$
\begin{aligned}
B_n'(f) &\leq 16(d!)^2 \sup_{(Y,Y',Z,Z')} \mathbb{E}\big[\overline{V}(Y_1\cap Y_2)\overline{V}(Y_1'\cap Y_3')\|\Delta_1 f(Z)\|^2\|\Delta_2 f(Z')\|^2\big] \\
&\leq \frac{16(d!)^2 C_2^{(5)}(d)}{n^4}.
\end{aligned}
$$

Thus, applying the estimate on $\mathbb{E}\|\Delta_1 f(X)\|^4$ from Lemma 6.5 and noting that $\mathbb{E}T = \Sigma_n$, we obtain from Lemma 2.4 that

$$
\begin{aligned}
\sqrt{\mathbb{E}\|\mathbb{E}[T|X] - \Sigma_n\|_{\text{H.S.}}^2} \\
\leq \Big(8\sqrt{d!C_4^{(2)}(d)} + 16d!\sqrt{C_2^{(5)}(d)} + 4\sqrt{C_4^{(1)}(d)}\Big)n^{-1/2} \\
\leq 1168(d!)^3(d+2)^2(2^{2^d+2d}d^{d/2})^2(3(R+1))^{4d}e^{8(2R+1)^d}n^{-1/2}.
\end{aligned}
$$
(6.24)

Moreover, notice that, by Theorem 3.1, matrix Σ_n is positive definite for large enough n and, for all $n > \overline{V}(K)e$,

$$\|\Sigma - \Sigma_n\|_{\text{H.S.}} \leq 116 \cdot 108^d (R+1)^{4d} e^{6\cdot9^d(R+1)^{2d}} (d!)^2(d+1)n^{-1/d}. \tag{6.25}$$

The result now follows from (6.21), (6.23)–(6.25) and Theorems 2.1 and 2.2.

SUPPLEMENTARY MATERIAL

Supplement (DOI: 10.1214/22-AAP1897SUPP; .pdf). Section A of the supplement contains the proof of Lemma 6.5. Section B contains proofs of Lemmas 6.6–6.9. In Section C, we present the proof of Theorem 4.1 and, in section D, the proof of Theorem 4.2.

REFERENCES

[1] BARBOUR, A. D. (1990). Stein's method for diffusion approximations. *Probab. Theory Related Fields* **84** 297–322. MR1035659 https://doi.org/10.1007/BF01197887
[2] BENTKUS, V. (2005). A Lyapunov type bound in R^d. *Theory Probab. Appl.* **49** 311–323. MR2144310 https://doi.org/10.1137/S0040585X97981123
[3] BHATTACHARYA, R. N. and RANGA RAO, R. (1976). *Normal Approximation and Asymptotic Expansions*. *Wiley Series in Probability and Mathematical Statistics*. Wiley, New York. MR0436272
[4] BICKEL, P. J. and BREIMAN, L. (1983). Sums of functions of nearest neighbor distances, moment bounds, limit theorems and a goodness of fit test. *Ann. Probab.* **11** 185–214. MR0682809
[5] BOUCHERON, S., LUGOSI, G. and MASSART, P. (2013). *Concentration Inequalities*: *A Nonasymptotic Theory of Independence*. Oxford Univ. Press, Oxford. With a foreword by Michel Ledoux. MR3185193 https://doi.org/10.1093/acprof:oso/9780199535255.001.0001
[6] CHATTERJEE, S. (2008). A new method of normal approximation. *Ann. Probab.* **36** 1584–1610. MR2435859 https://doi.org/10.1214/07-AOP370
[7] CHATTERJEE, S., DIACONIS, P. and MECKES, E. (2005). Exchangeable pairs and Poisson approximation. *Probab. Surv.* **2** 64–106. MR2121796 https://doi.org/10.1214/154957805100000096
[8] CHATTERJEE, S. and MECKES, E. (2008). Multivariate normal approximation using exchangeable pairs. *ALEA Lat. Am. J. Probab. Math. Stat.* **4** 257–283. MR2453473
[9] CHATTERJEE, S. and SEN, S. (2017). Minimal spanning trees and Stein's method. *Ann. Appl. Probab.* **27** 1588–1645. MR3678480 https://doi.org/10.1214/16-AAP1239
[10] CHEN, L. H. Y., GOLDSTEIN, L. and SHAO, Q.-M. (2011). *Normal Approximation by Stein's Method*. *Probability and Its Applications* (*New York*). Springer, Heidelberg. MR2732624 https://doi.org/10.1007/978-3-642-15007-4
[11] DÖBLER, C. and PECCATI, G. (2017). Quantitative de Jong theorems in any dimension. *Electron. J. Probab.* **22** Paper No. 2, 35. MR3613695 https://doi.org/10.1214/16-EJP19
[12] DUERINCKX, M. (2021). On the size of chaos via Glauber calculus in the classical mean-field dynamics. *Comm. Math. Phys.* **382** 613–653. MR4223483 https://doi.org/10.1007/s00220-021-03978-3
[13] DUNG, N. T. (2019). Explicit rates of convergence in the multivariate CLT for nonlinear statistics. *Acta Math. Hungar.* **158** 173–201. MR3950207 https://doi.org/10.1007/s10474-019-00917-6
[14] FANG, X. and KOIKE, Y. (2020). Large-dimensional central limit theorem with fourth-moment error bounds on convex sets and balls. ArXiv preprint. Available at arXiv:2009.00339.
[15] FANG, X. and KOIKE, Y. (2021). High-dimensional central limit theorems by Stein's method. *Ann. Appl. Probab.* **31** 1660–1686. MR4312842 https://doi.org/10.1214/20-aap1629
[16] FANG, X. and KOIKE, Y. (2022). New error bounds in multivariate normal approximations via exchangeable pairs with applications to Wishart matrices and fourth moment theorems. *Ann. Appl. Probab.* **32** 602–631. MR4386537 https://doi.org/10.1214/21-aap1690
[17] GAUNT, R. E. (2016). Rates of convergence in normal approximation under moment conditions via new bounds on solutions of the Stein equation. *J. Theoret. Probab.* **29** 231–247. MR3463084 https://doi.org/10.1007/s10959-014-0562-z
[18] GLORIA, A. and NOLEN, J. (2016). A quantitative central limit theorem for the effective conductance on the discrete torus. *Comm. Pure Appl. Math.* **69** 2304–2348. MR3570480 https://doi.org/10.1002/cpa.21614
[19] GOLDSTEIN, L. and PENROSE, M. D. (2010). Normal approximation for coverage models over binomial point processes. *Ann. Appl. Probab.* **20** 696–721. MR2650046 https://doi.org/10.1214/09-AAP634
[20] GÖTZE, F. (1991). On the rate of convergence in the multivariate CLT. *Ann. Probab.* **19** 724–739. MR1106283

[21] HALL, P. (1988). *Introduction to the Theory of Coverage Processes*. *Wiley Series in Probability and Mathematical Statistics*: *Probability and Mathematical Statistics*. Wiley, New York. MR0973404 https://doi.org/10.1016/0167-0115(88)90159-0

[22] HUG, D., LAST, G. and SCHULTE, M. (2016). Second-order properties and central limit theorems for geometric functionals of Boolean models. *Ann. Appl. Probab.* **26** 73–135. MR3449314 https://doi.org/10.1214/14-AAP1086

[23] KASPRZAK, M. J. and PECCATI, G. (2023). Supplement to "Vector-valued statistics of binomial processes: Berry–Esseen bounds in the convex distance." https://doi.org/10.1214/22-AAP1897SUPP

[24] LACHIÈZE-REY, R. and PECCATI, G. (2017). New Berry–Esseen bounds for functionals of binomial point processes. *Ann. Appl. Probab.* **27** 1992–2031. MR3693518 https://doi.org/10.1214/16-AAP1218

[25] LACHIÈZE-REY, R., PECCATI, G. and YANG, X. (2022). Quantitative two-scale stabilization on the Poisson space. *Ann. Appl. Probab.* **32** 3085–3145. MR4474528 https://doi.org/10.1214/21-aap1768

[26] LAST, G. and PENROSE, M. (2018). *Lectures on the Poisson Process*. *Institute of Mathematical Statistics Textbooks* **7**. Cambridge Univ. Press, Cambridge. MR3791470

[27] LEVINA, E. and BICKEL, P. J. (2005). Maximum likelihood estimation of intrinsic dimension. In *Advances in NIPS* (K. L. Saul, Y. Weiss and L. Bottou, eds.) **17**.

[28] LOTZ, M., MCCOY, M. B., NOURDIN, I., PECCATI, G. and TROPP, J. A. (2020). Concentration of the intrinsic volumes of a convex body. In *Geometric Aspects of Functional Analysis. Vol. II. Lecture Notes in Math.* **2266** 139–167. Springer, Cham. MR4175761 https://doi.org/10.1007/978-3-030-46762-3_6

[29] MECKES, E. (2009). On Stein's method for multivariate normal approximation. In *High Dimensional Probability V*: *The Luminy Volume. Inst. Math. Stat.* (*IMS*) *Collect.* **5** 153–178. IMS, Beachwood, OH. MR2797946 https://doi.org/10.1214/09-IMSCOLL511

[30] MERIKOSKI, J. K. and VIRTANEN, A. (1997). Bounds for eigenvalues using the trace and determinant. *Linear Algebra Appl.* **264** 101–108. MR1465858 https://doi.org/10.1016/S0024-3795(97)00067-0

[31] MORAN, P. A. P. (1958). Random processes in genetics. *Proc. Camb. Philos. Soc.* **54** 60–71. MR0127989 https://doi.org/10.1017/s0305004100033193

[32] NAZAROV, F. (2003). On the maximal perimeter of a convex set in $\mathbb{R}^n$ with respect to a Gaussian measure. In *Geometric Aspects of Functional Analysis.* (V. D. Milman and G. Schechtman, eds.). *Lecture Notes in Math.* **1807** 169–187. Springer, Berlin. MR2083397 https://doi.org/10.1007/978-3-540-36428-3_15

[33] NOURDIN, I. and PECCATI, G. (2012). *Normal Approximations with Malliavin Calculus*: *From Stein's Method to Universality. Cambridge Tracts in Mathematics* **192**. Cambridge Univ. Press, Cambridge. MR2962301 https://doi.org/10.1017/CBO9781139084659

[34] NOURDIN, I., PECCATI, G. and RÉVEILLAC, A. (2010). Multivariate normal approximation using Stein's method and Malliavin calculus. *Ann. Inst. Henri Poincaré Probab. Stat.* **46** 45–58. MR2641769 https://doi.org/10.1214/08-AIHP308

[35] NOURDIN, I., PECCATI, G. and YANG, X. (2022). Multivariate normal approximation on the Wiener space: New bounds in the convex distance. *J. Theoret. Probab.* **35** 2020–2037. MR4488569 https://doi.org/10.1007/s10959-021-01112-6

[36] PENROSE, M. (2003). *Random Geometric Graphs. Oxford Studies in Probability* **5**. Oxford Univ. Press, Oxford. MR1986198 https://doi.org/10.1093/acprof:oso/9780198506263.001.0001

[37] PENROSE, M. D. and YUKICH, J. E. (2001). Central limit theorems for some graphs in computational geometry. *Ann. Appl. Probab.* **11** 1005–1041. MR1878288 https://doi.org/10.1214/aoap/1015345393

[38] RAIČ, M. (2019). A multivariate Berry–Esseen theorem with explicit constants. *Bernoulli* **25** 2824–2853. MR4003566 https://doi.org/10.3150/18-BEJ1072

[39] REINERT, G. and RÖLLIN, A. (2009). Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition. *Ann. Probab.* **37** 2150–2173. MR2573554 https://doi.org/10.1214/09-AOP467

[40] REINERT, G. and RÖLLIN, A. (2010). Random subgraph counts and U-statistics: Multivariate normal approximation via exchangeable pairs and embedding. *J. Appl. Probab.* **47** 378–393. MR2668495 https://doi.org/10.1239/jap/1276784898

[41] RINOTT, Y. and ROTAR, V. (1996). A multivariate CLT for local dependence with $n^{-1/2} \log n$ rate and applications to multivariate graph related statistics. *J. Multivariate Anal.* **56** 333–350. MR1379533 https://doi.org/10.1006/jmva.1996.0017

[42] RINOTT, Y. and ROTAR, V. (1997). On coupling constructions and rates in the CLT for dependent summands with applications to the antivoter model and weighted U-statistics. *Ann. Appl. Probab.* **7** 1080–1105. MR1484798 https://doi.org/10.1214/aoap/1043862425

[43] SANTALÓ, L. A. (2004). *Integral Geometry and Geometric Probability*, 2nd ed. *Cambridge Mathematical Library*. Cambridge Univ. Press, Cambridge. With a foreword by Mark Kac. MR2162874 https://doi.org/10.1017/CBO9780511617331

[44] SCHNEIDER, R. and WEIL, W. (2008). *Stochastic and Integral Geometry. Probability and Its Applications* (*New York*). Springer, Berlin. MR2455326 https://doi.org/10.1007/978-3-540-78859-1

[45] SCHULTE, M. and YUKICH, J. E. (2019). Multivariate second order Poincaré inequalities for Poisson functionals. *Electron. J. Probab.* **24** Paper No. 130, 42. MR4040990 https://doi.org/10.1214/19-ejp386

[46] SCHULTE, M. and YUKICH, J. E. (2021). Rates of multivariate normal approximation for statistics in geometric probability. ArXiv preprint.

[47] SHAO, Q.-M. and SU, Z.-G. (2006). The Berry–Esseen bound for character ratios. *Proc. Amer. Math. Soc.* **134** 2153–2159. MR2215787 https://doi.org/10.1090/S0002-9939-05-08177-3

[48] STEELE, J. M. (1986). An Efron–Stein inequality for nonsymmetric statistics. *Ann. Statist.* **14** 753–758. MR0840528 https://doi.org/10.1214/aos/1176349952